PSYCHOLOGICAL REVIEW PUBLICATIONS Whole No. 169
1913

PRINCETON LIBRARY

Psychological Monographs

EDITED BY
JAMES ROWLAND ANGELL, UNIVERSITY OF CHICAGO
H. C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, JOHNS HOPKINS UNIVERSITY (*J. of Exp. Psychol.*)
SHEPHERD I. FRANZ, GOVT. HOSP. FOR INSANE (*Bulletin*) and
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*Index*)

UNIVERSITY OF IOWA STUDIES IN PSYCHOLOGY

No. VII

Edited
BY
CARL E. SEASHORE

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.
AND LANCASTER, PA.

AGENTS: G. E. STECHERT & CO., LONDON (2 Star Yard, Carey St., W. C.)
LEIPZIG (Hospitalstrasse, No. 27); PARIS (76 rue de Rennes)

DESCRIPTION OF AN UNUSUAL CASE OF PARTIAL COLOR BLINDNESS

BY

MABEL CLARE WILLIAMS

CONTENTS

Introduction
Color matching
Color recognition
Color equations
Color sensitivity
Color discrimination
Color fields
Relative stimulating values of the different wave lengths
Spectral limits
Threshold differences for wave-length
Matches between mixtures of two homogeneous lights and a single homogeneous light
The achromatic band
Miscellaneous

Introduction

The subject of this study will be referred to as J. H. He was a student in the University of Iowa during the year 1914-1915. His matriculation card gives as the place and date of his birth: Edinburgh, Scotland, April 27, 1890. He is of medium height and weight, in good physical health, and shows no evidence of disease or condition that would in any way influence his vision. He does not use tobacco, alcohol, or other drugs. His eyes are light blue in color with slightly bluish sclera, well set and wholly "normal" in appearance. For constant use he wears spectacles with peritoric lenses, spheres +1.25 dioptres over each eye, fitted in London without cycloplegic. Dr. Heard, physician in charge of the ophthalmic clinic in the College of Medicine in the University of Iowa, re-examined J. H.'s eyes, also without cycloplegic, and reports "vision right eye 6/3, left eye 6/3, showing him to be a hyperope since he reads at 6

meters what he should read at 3, and he reads this equally well with the +1.25 spheres. His accommodation is evidently very active. Since he refused a cycloplegic, this is only a measure of his manifest, not his latent, hyperopia. Ophthalmic examination reveals no astigmatic error. Fundi those of the typical hyperope. Retina granular in appearance due to pigment. Arteries not tortuous. Media clear."

Little family history of a satisfactory sort is available. His father, who died when J. H. was fifteen years old, was known to be color blind (letter from his widow, the mother of J. H. and herself not color blind) but there is nothing to indicate the type. Of the three brothers of J. H., two had probably normal color vision and in general appearance are said to have resembled the mother; while the third brother, A. H., and J. H. himself, resembled strongly the (color blind) father and were themselves color blind. The brother A. H., recently dead, apparently exhibited the same sort of color defect as J. H. As boys and young men together they often compared colors of objects and agreed always with each other but not with the other two brothers, only one of whom is now living. The two sisters are not color blind so far as J. H. knows. J. H. is the only member of the family who has been in this country.

J. H. has at all times been a most satisfactory observer. He was very careful and painstaking and unusually accurate, due in part perhaps to extensive training in biology and chemistry. He has done much microscopic work. Considering the fact that he has never been shown or told any of the results of the experiments even to the present time, his coöperation and willingness are all the more praiseworthy. It is therefore most pleasing to the writer to make public acknowledgement of indebtedness to him and to thank him. He began the experiments with the pre-established conviction that he was blind to blue and yellow, and was allowed to continue in that belief without discussion.

Attention was first called to the case in March 1915 through a mutual friend, Mr. C. J. Knock, fellow in the department of psychology at Iowa. It was noticed that only two color names

were used, "red" and "green", and that "red" was applied indifferently to reds and yellows and "green" to greens and blues. There has never been any noteworthy variation from this throughout the whole period of experimentation. In connection with hospital work, he once had to diagnose with the attending physician a case of "yellow jaundice". The patient looked "red" to him and very queer. He thought she must be an American Indian, but the features were not right and she had light hair. "Her skin and eyes were red and she was a funny looking woman."

Experimental work of a preliminary sort was begun in the Iowa Psychological Laboratory on March 5th and continued at frequent intervals until June, when J. H. had suddenly to leave. Fortunately he was again available for study two weeks in December, 1915. Since the resources of the Iowa Laboratories had been practically exhausted and it was now advisable to work with spectral light, the writer went to Chicago to canvass the situation there. Professor Carr at that time suggested that the best opportunity for continuing the work was to be found in the Nela Research Laboratory of the National Lamp Works of the General Electric Company at Cleveland, Ohio. Accordingly, invitation was sought from the Director of the Laboratory, Dr. E. P. Hyde, and most cordially extended by him. The writer went to Cleveland early in December to prepare in advance for J. H. who arrived from Missouri on December 19th and remained until the night of January 1st, 1916. Traveling expenses for both were generously provided from the research funds of the Iowa Graduate College.

At the Nela Research Laboratory little could have been accomplished in so short a time without the considerate aid of Dr. H. M. Johnson and Mr. M. Luckiesh, to both of whom the writer is most deeply indebted. Dr. Cobb, Dr. Troland and other members of the staff contributed in less extensive but very important ways. Every facility possessed by the Laboratory was lent unreservedly. A valuable instrument (Hilger constant deviation wave-length spectrometer) was borrowed

from Dr. Dunlap of Johns Hopkins and filled an immediate need.

The report of this case, though apparently extensive, should be regarded as a preliminary survey; and it is hoped that at some future time the study may be continued at the Nela Research Laboratory. The color names as used by J. H. have been quite disregarded. Control observers, usually three in number, were used in all tests where relative values only were obtainable. For the most part the several experiments are reported in their chronological order. The tests have all been conducted by the writer and are reliable to the extent claimed.

Color matching. (*Colored worsteds*)

A large well-mixed assortment of worsted skeins of many colors was spread out in good day light, on a table, over a large square of grey cloth. These skeins were not numbered and will be referred to by color only. No color names were used by examiner or subject.

(A) *Red* standard. A rich fairly dark red skein was handed to J. H. and he was asked to select from the pile all those skeins which resembled it in color but not necessarily in brightness. He selected out all reds and yellows, some sixteen skeins in all. He was then asked to re-sort these selected skeins according to two samples, the original red and a good yellow which he had included as a match for the red. He tried this, but gave it up as impossible.

(B) *Green* standard. Following the same procedure he selected to match the green sample thirty-two blues and greens of varying brightness and saturation. No red or reddish skeins were chosen. Given two samples, green and blue, and asked to re-sort his selections into two piles he again failed.

(C) *Yellow* standard. J. H. selected to match the yellow skein, the same skeins which he had sorted out to match the red and which had been returned to the main assortment. He added however a few more reds and yellows.

(D) *Blue* standard. The matches for the blue were identical with those for the green sample.

(E) *Grey.* Without sample, he was asked to select some "greys". Fifteen skeins were chosen, five of which were nearly true greys, eight were tans, and two were brown.

Color recognition. (*Nagel cards. Fifth edition*)

The available set of cards was old and somewhat soiled and perhaps faded. The results of the test should not be too rigidly interpreted. In the first four experiments (A, B, C, D) the regular procedure was followed.

(A) Select the card with red or reddish dots: A 12, A 6, A 1 were selected in the order given.

(B) Only red spots: A 3, and possibly A 15. He was not sure about A 15.

(C) Only green spots: A 5, A 13. He called the greyish spots on A 13 "green."

(D) Only grey spots: A 6. He also accepted A 9 after his attention was called to it and indicated the true grey spots on A 6 when asked to do so.

(E) For J. H., card B 1 consists of "plain red" (the shades of yellow) and "dark red or brownish".

(F) On card B 2 all the spots are "red, but the bright ones are redder; the others are brownish".

(G) Point out all the A cards with red spots: 6, 15, 12, 14, 3, 1, 11 (the yellow spots), 7, 16, 8, 10. Of the remaining A cards, "5 is green, 4 has green spots (correctly indicated), 13 is green with the greys showing a trace of green; 9 has shades of very dark green."

Color equations. (*Color mixing by rotation*)

Disks of two sizes, 5.8 and 9.6 cm. radii, were cut from Hering colored papers: red (old No. 1), green (old No. 7, new No. 9), blue (old No. 10, new No. 13), yellow (old No. 5, new No. 5), and black and white. Care was taken to guard against fatigue and an exact match was insisted upon. A small disk of single color was mounted over large disks in combination and the following equations were secured.

(A) Red 360° = yellow 270° + black 90°

(B) Green 360° = blue 226° + white 134°
(C) Blue 360° = green 132° + white 21° + black 207°
(D) Yellow 360° = red 257° + white 45° + black 58°

The four colors were then combined in pairs to match grey (black and white).

(E) Blue 279° + red 81° = white 51° + black 309°
(F) Green 222° + red 138° = white 111° + black 249°
(G) Yellow 173° + blue 187° = white 150° + black 210°
(H) Yellow 167° + green 193° = white 129° + black 231°

In the three experiments thus far described, the results of the Nagel test seem not to be very decisive. In the conditions under which they were conducted, their chief significance is in indicating that some defect in color vision is obviously present. Too much stress is placed upon the color names which J. H. can apply successfully to red and green. The worsted test is suggestive and quite corroborated by the color equation tests. J. H. evidently has his color system reduced to two processes: that which conditions what he calls "red" and applies consistently to both red and yellow; and that which conditions what he calls "green" and applies consistently to green and blue. Real red and green are not confused nor are yellow and blue. These color processes are apparently physiologically antagonistic since a match for grey may be produced by proper combinations on the color wheel. The darkest grey is matched by blue (called "green") and red (called "red"); the next lighter grey is matched by green ("green") and red ("red"); the next lighter by yellow ("red") and green ("green"); and the lightest grey by yellow ("red") and blue ("green"). One would naturally expect the combination of yellow and green to require the most luminous grey as its equivalent. This discrepancy is probably in part due to the fact that only one trial was made and, in spite of frequent rests, fatigue may have entered, as all four of the grey equivalents were made at one rather tedious sitting and the yellow-green combination was tried last. Subsequent experiments (luminousity curves) show that for day-light adaptation, green is much nearer to spectral yellow in luminousity than is blue.

Color sensitivity. (*Colored rings on white rotating disks*)

This is essentially the same test described by Hayes, (1). The disks, 9.6 cm. in radius, were somewhat larger than those made by Hayes. They were cut from white bond paper, several thicknesses of which were used to avoid transparency. The rings of color (Hering papers red, green, yellow and blue as previously described and grey No. 9 from the Hering 50 greys) were 1 cm. wide, and were pasted smoothly midway between the center and the periphery of the disks. It was not possible to keep all the four colors and grey on the wheel at once[1] so each color was paired with the same grey and the observer was asked to decide whether a given color or grey was exposed. More foreknowledge on the part of the observer was implied here than in Hayes' tests and the threshold values are probably on that account lowered. Chance determined whether a portion of a colored ring or of the grey should be exposed. Records were obtained also from three control observers, Dr. R. H. Sylvester, Mr. C. J. Knock, and Mr. P. J. Sodergren, all members of the psychology department at Iowa. The results which represent the minimal exposure of color necessary for recognition, are stated in degrees. Ten determinations were made on each color paired with grey for J. H. and five for the controls. The averages and mean variations are presented in Table I.

TABLE I

	J.H.		R.H.S.		C.J.K.		P.J.S.	
	Ave.	m. v.	Ave.	m. v.	Ave.	m. v.	Ave.	m. v.
Red or	11.2°	1.8°	2.7°	0.2°	2.0°	0.2°	3.0°	0.8°
Grey	12.0	0.8	2.6	0.7	4.0	0.6	3.4	0.5
Green	9.1	0.9	2.5	0.0	3.2	1.4	3.6	0.5
Grey	11.2	1.6	2.7	0.2	2.1	0.7	3.6	0.5
Blue	12.3	2.5	23.0	5.6	2.3	0.4	6.2	1.8
Grey	18.3	5.0	6.8	2.7	4.0	1.6	4.6	1.6
Yellow	36.2	4.2	4.2	0.9	4.1	1.4	5.6	1.3
Grey	14.0	1.4	4.0	0.4	3.7	1.2	4.1	1.1

[1] This was actually done for J. H. however, and when he did not know what to expect, red was recognized at 13°, green at 21°, yellow ("red") at 31°, and blue ("green") at 33°. Grey was called "dirty white" at 10°.

The results of this test, which should be regarded as having relative value only, indicate a lesser degree of sensitiveness for the color stimuli in the case of J. H., as compared with the three control observers. The only exception to this is with the blue for R. H. S., whose readings for this color scattered widely and were very uncertain.

Color discrimination. (*The Lovibond Tintometer, 2*)

This instrument offers a ready means for measuring color sensitivity and discrimination when pure color stimuli are not demanded. The colored glass ray filters which are provided with the instrument are carefully standardized according to an arbitrary scale, and the units are recoverable, convenient and fairly permanent. The filters are not even approximately monochromatic, however, although to the eye the transmitted hues appear to be fairly pure.

The tintometer was set up where north daylight fell directly upon its windows. A ground-glass strip was placed between the slips and the source of light as an aid in securing a homogeneous surface. It was necessary to compound the glass slips in order to secure the proper increments from the available assortment. Four or even five slips could be bound together quickly with a rubber band and handled as easily as one. Each slip doubtless cut down the intensity of the transmitted light, for even clear glass will do this to the extent of 8%. A clear or white glass was always used in connection with the standard to help equalize this effect. Lateral light was excluded from the observers' eyes by means of the hood of the instrument. The colors were viewed with both eyes open.

A simple form of the method of constant stimuli was found satisfactory. Two slips, say red 10, were placed in two adjacent windows of the instrument, then, after a few minutes of skirmishing, the increment was found which when added to one of the windows containing red 10, made it just perceptibly "redder" than the other. A difference which yielded 80% of correct judgments was regarded as the approximate threshold. This increment was then fixed to one of the red standard slips

and one hundred comparison judgments were made. The right and left positions of the two stimuli were determined in accordance with a chance key already prepared. The observer either closed his eyes or looked away while the slips were being shifted and always pronounced judgment quickly without staring at the colors. From the number of correct judgments and the size of the increment, the discrimination threshold which probably would yield 75% of correct judgments was computed by means of the well known Fullerton-Cattell table.

In Table II are given: the initials of the observers, J. H. and three controls; the composition of the two stimuli compared, St. 1 and St. 2; the threshold increment in terms of arbitrary tintometer units; the percentage of correct judgments in one hundred trials; and the probable threshold for each of the four colors. It should be remembered that J. H. identifies yellow with red and blue with green.

TABLE II

Observer	St. 1	St. 2	Incr.	%	75%
		RED			
J.H.	R 10 + clear	R 10 + 5	5.0	85	3.20
R.H.S.	R 10 + clear	R 10 + .2 + .1	.3	83	.21
C.J.K.	R 10 + clear	R 10 + .1 + .1	.2	83	.14
P.J.S.	R 10 + clear	R 10 + .1	.1	83	.07
		GREEN			
J.H.	B 10 + Y 10 + B5 + Y5	B10 + Y10 + B10 + Y10	5.0	72	5.80
R.H.S.	B 10 + Y 10	B10 + Y10 + B1 + Y1	1.0	89	.60
C.J.K.	B 10 + Y 10 + clear	B10 + Y10 + B.3 + Y.3	.3	88	.17
P.J.S.	B 10 + Y 10 + clear	B10 + Y10 + B.2 + Y.2	.2	83	.14
		BLUE			
J.H.	B 10 + clear	B 10 + 1 + 1 + 1	3.0	75	3.00
R.H.S.	B 10 + clear	B 10 + .5	.5	81	.39
C.J.K.	B 10 + clear	B 10 + .5	.5	86	.30
P.J.S.	B 10 + clear	B 10 + .1	.1	76	.10
		YELLOW			
J.H.	Y 10 + 1 + 1	Y 10 + 5 + 1	4.0	79	3.30
R.H.S.	Y 10 + clear	Y 10 + .3 + .2	.5	81	.39
C.J.K.	Y 10 + clear	Y 10 + .1 + .1	.2	64	.37
P.J.S.	Y 10 + clear	Y 10 + .3 + .2	.5	87	.30

In the case of J. H., St. 1 is slightly different from that of the control observers for the colors yellow and green. An attempt was made to use for him the yellow which could not be distinguished from red 10, which was yellow 10 + 1 + 1. Green

was somewhat difficult to secure, since no green slips were available, and blue and yellow had to be combined for it. However, a very good green resulted from this combination; and the green which matched blue 10 was compounded from Y 10 + B 10 + Y 5 + B 5 + Y 1 + B 1. These six slips were too thick when bound together to fit the receiving stalls of the tintometer, so the last two were omitted, and the four designated in the table retained.

These tests with the tintometer indicate that J. H. has a decidedly higher discrimination threshold than the normal observers. It is however questionable whether he is really judging depth of color at all; the probability is that some other factor, perhaps luminosity, served as a criterion. To what extent this is also true of the other observers is uncertain; they are at any rate, more sensitive to color than J. H. and, under identical conditions, make better discrimination records. The results, even as they stand, are too favorable for J. H., (see his hue discrimination records, p. 20 f).

Color fields

A recent model of the Meyerowitz self-recording perimeter was secured and fitted with new stimulus patches 9 cm. square, cut from the four Hering colored papers and white. The self-recording device was discarded and the records read in degrees directly from the arms. The color carriers were moved at such a rate that approximately ten seconds were required for one determination. The unused eye was comfortably bandaged. A color, not known to the observer, was moved inward from the periphery and the position noted where it first was named correctly. Two determinations for each of the five stimulus patches were made in each of the four cardinal meridians for each eye, with J. H. and the three regular control observers. J. H., of course, reported only two colors and white. His records again have relative value only, and achieve their significance when compared with those made upon observers with normal[2] color

[2] C.J.K. has complete central scotoma in his right eye and perhaps slight color weakness.

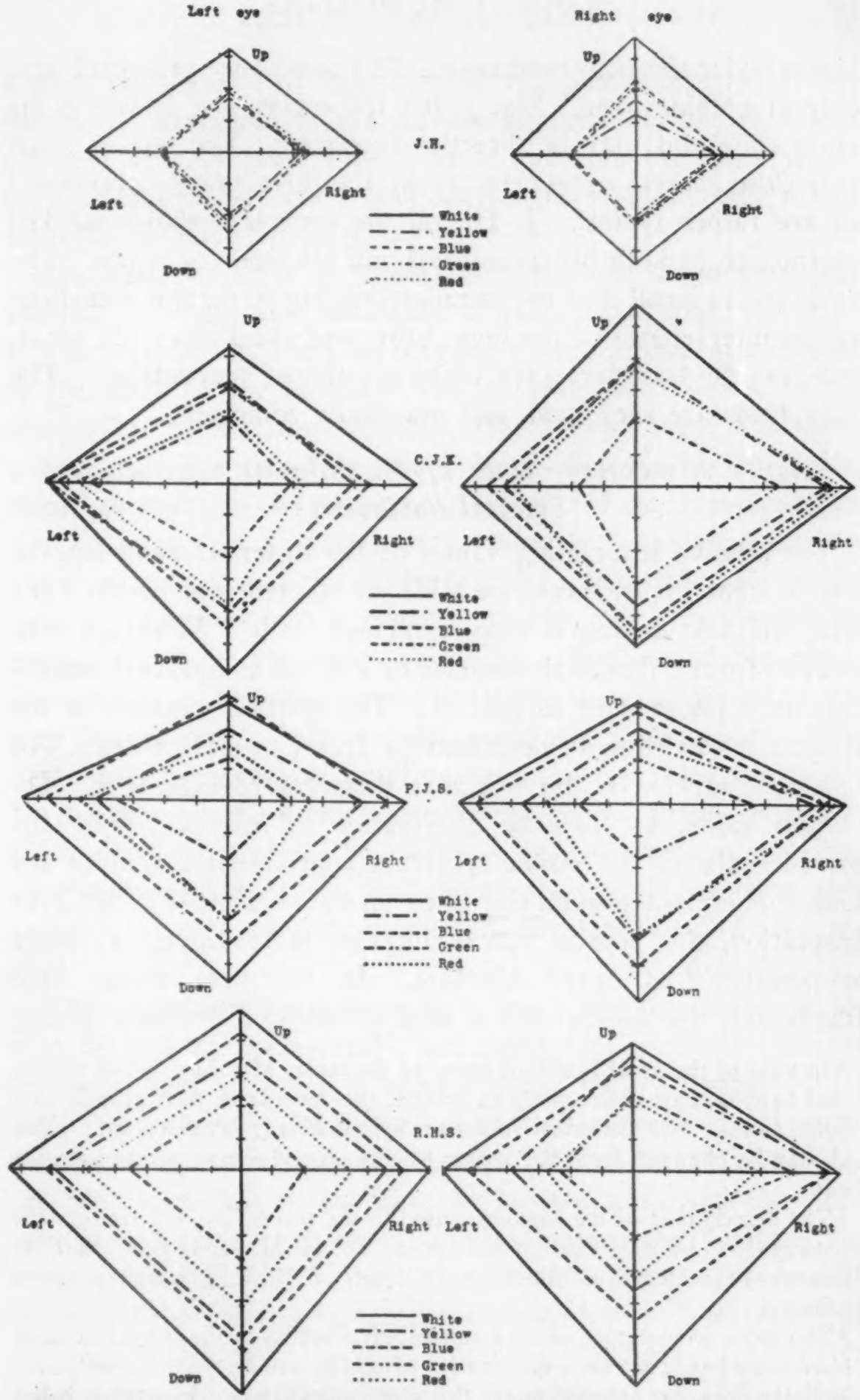

FIG. 1. Visual fields of J.H. and three control observers. Drawn to same scale. Units on axes = 10°.

vision under identical conditions. The results are presented here only in graphic form,[3] Fig. 1, but the curves are drawn to the same scale and may be directly compared. The curves show very close general agreement among the three control observers; all are rather typical. J. H., on the contrary, shows marked contraction of both his color fields and his field for white. The fields are so small that the familiar limiting in certain meridians by the interference of the nose, brow and cheek does not occur, and even the boundaries for white are almost symmetrical. The color fields are very small and practically coincident.

Relative stimulating values of the different wave lengths[4] (*Spectral luminosity*)

The relative stimulating values of the different wave lengths were sought for by means of a Hilger spectroscope of the fixed arm, constant deviation type, combined with a Whitman disk so as to form a flicker photometer or a direct comparison heterochromatic photometer as desired. The essential feature of the apparatus was recently described by Ives (3) and others. The reflecting surface of the Whitman disk was covered with magnesium oxide, so that the distribution of energy among the wave-lengths of the visible spectrum approximated that of the source of illumination on the disk. In the tests made under light adaptation, this source was a tungsten lamp having a "color temperature"[5] of 2327° Absolute. In the tests under dark adaptation, the source was a similar but smaller lamp giving

[3] In view of the great length of many of the tables of values in this report, it has been deemed unnecessary to present the records in both graphic and tabular forms. The numerical values are systematically tabulated and on file and may be obtained from the writer by any one who may perchance wish them.

[4] The descriptions of the apparatus used in the succeeding experiments are substantially those very kindly furnished by Dr. H. M. Johnson of the Nela Research Laboratory, at which place most of these experiments were conducted.

[5] The color temperature of an incandescent filament is the temperature of a black body giving the same distribution of energy among the different wavelengths as does the filament under the given conditions. This datum being specified, the spectral character of the radiation emitted by the source may be readily ascertained from well known formulae.

a color temperature of 2355° Absolute. The brightness of the disk was determined by the position of the lamp and was kept constant at 32.8 candles per square meter in the work under light adaptation, and at 0.086 candles per square meter in the work under dark adaptation.

The observed field, as limited by a suitable diaphragm, was circular in form, and subtended a solid visual angle of 2.4°. It was observed through an opening 1 mm. in diameter, placed in the plane of the image of the collimator-slit formed in the minimally deviated wave-length. The observer's eye was about 1 cm. from the aperture, which is as close as his eye-lashes would permit. In the tests by the method of direct comparison, the Whitman disk was stationary, and turned so that one of its radial edges bisected the field vertically. The left half of the field as observed was a portion of the disk, illuminated by tungsten light. The right half was a portion of the prism illuminated in monochromatic light, the wave-length of which was determined by the angular position of the prism. The limiting diaphragm was in all cases completely filled with light. In the tests by the flicker method, the entire circular field was illuminated successively by colorless and monochromatic light. The surroundings of the field were dark except for light scattered in the optical system and in the refractive media of the eye.

The source of homogeneous light was a tungsten lamp. In the tests by both the flicker and direct comparison methods, the spectrometer lamp occupied a fixed position with reference to the collimator-slit. In the direct comparison method, the observer sought an equation of the brightness of the colored portion of the field with that of the disk, by himself adjusting the width of the collimator-slit, controlled by a micrometer screw, with scale reading directly to the 0.01 mm. The distance between the lamp and the slit was increased to secure lower illumination for the tests under dark adaptation and a ground-glass surface was placed between the filament and the collimator-slit, to give uniformity of brightness over the colored portion of the field and further to reduce illumination.

In the experiments according to the flicker method, the speed of the disk was controlled, within limits, by means of resistance in the motor circuit.

The entire system was covered with black felt, and the eyes of the observer were protected from stray light from the lamps. Day-light was excluded from the room. In the tests under light adaptation, the general illumination of the room was from four 150 watt "Mazda C-2" lamps in an overhead fixture. These lamps have tungsten filaments in a gas-filled bulb, the selective absorption of which is such as to transmit light approximating that of afternoon sunlight in spectral character. In the work under dark adaptation, the general illumination was eliminated and fair dark-room conditions were obtained. The observer remained in the darkened room about twenty minutes before beginning the experiments and kept his eyes closed except during the short interval when making the comparisons.

The flicker settings under dark adaptation became very uncertain except for the middle regions of the spectrum, perhaps because the speed of the motor was not variable within sufficiently wide limits.

In making the settings, a definite procedure was followed. The observer himself turned the micrometer which varied the slit-width which controlled the amount of light admitted to the prism of the spectrometer. The scale readings were all made and recorded by the experimenter. Beginning at 720 $\mu\mu$, five settings were made at regular intervals of 20 $\mu\mu$ until 420 $\mu\mu$ was reached. Then the spectrum was similarly surveyed in the reverse order. Of the ten readings thus secured for each wave-length selected, five were made with the luminosity varying from less than that of the colorless comparison field, to equality, and the other five varied from greater luminosity to apparent equality. These two directions of approach alternated regularly in both the direct comparison and the flicker methods, which themselves alternated, five direct comparison readings being made, then five flicker readings. Ten readings, taken in the double fatigue order, were thus secured for each 20 $\mu\mu$ interval

of the spectrum within the limits indicated, for both the direct comparison and flicker methods. With otherwise identical procedure, readings were obtained for both photopic and scotopic vision for each eye of J. H. separately, and for the right eye of three control observers, Dr. L. T. Troland, Mr. Everett Schmidt (a graduate of one of the Cleveland technical high schools), and from the writer.

(A). *Photopic luminosity.* (*Daylight adaptation*)

The records are given in graphic form in Figs. 2, 3, 4, 5 and 6, and represent the reciprocals of the slit-widths in mm. plotted against the ordinates, and the wave-lengths as abscissae. The result is a luminosity curve for a prismatic spectrum of tungsten at a temperature of 2385° Absolute.

J. H., for whom such work was new, had some difficulty at first with the direct comparison method and the results for this method never became so constant as for the flicker method. The mean variations (not presented) show this very clearly. His curves for the flicker method bear a close general resemblance to those of the control observers and also those which have been published by Ives and other investigators. In his direct comparison curves, there is a noticeable break at 560 $\mu\mu$; here the slit was opened much wider than for the wave-lengths immediately preceding and following, 540 $\mu\mu$ and 580 $\mu\mu$ respectively. This is probably due to the fact that there is an achromatic or neutral band in this region of his spectrum, near 560 $\mu\mu$, and he apparently opens the slit wider in the effort to admit color. For the flicker method this does not occur, as the whole field then becomes homogeneous and he is not directly conscious of equating two separate fields.

J. H.'s descriptive comments about the various colored lights are interesting and are given here in his own wording as they were recorded during the progress of the experiment.

720 $\mu\mu$: "Deep red."
700 $\mu\mu$: "Red; good deep bright red."
680 $\mu\mu$: "Nice bright red."
660 $\mu\mu$: "Good bright red."

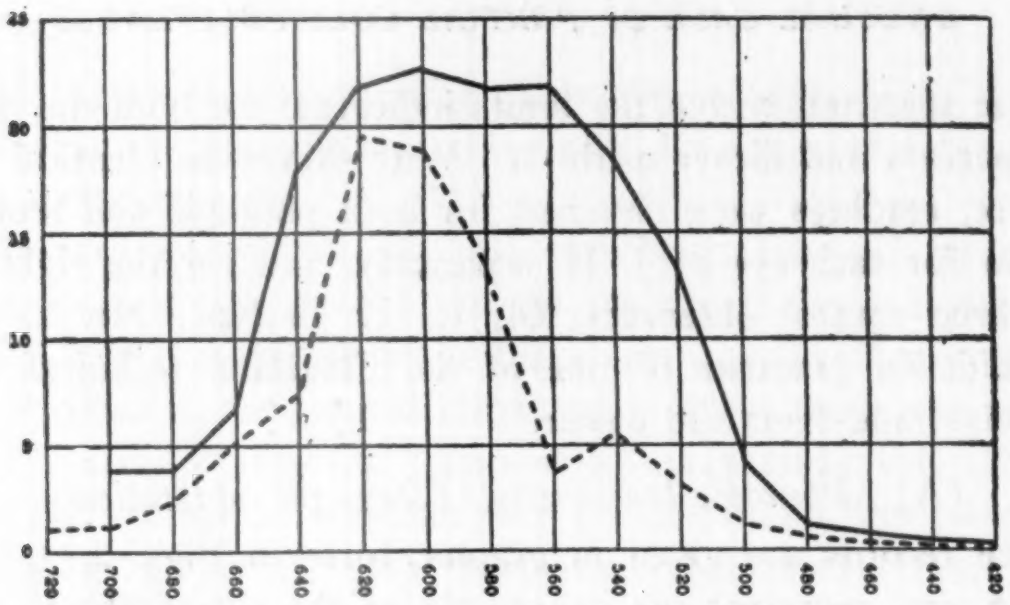

FIG. 2. Photopic luminosity curve for J. H. Right eye. - - - - - = direct comparison, ——— = flicker method.

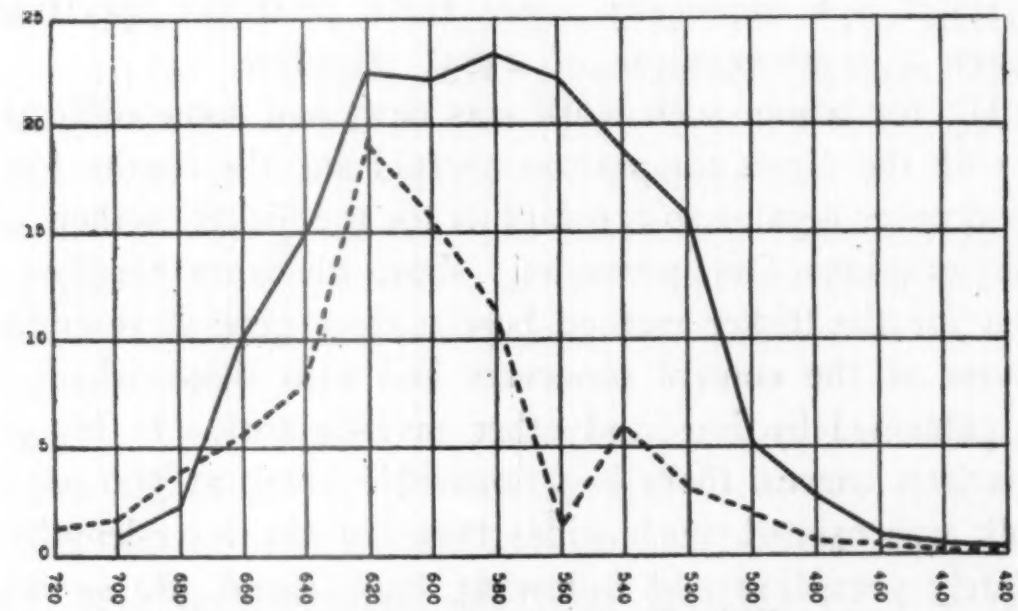

FIG. 3. Photopic luminosity curve for J. H. Left eye. - - - - - = direct comparison, ——— = flicker method.

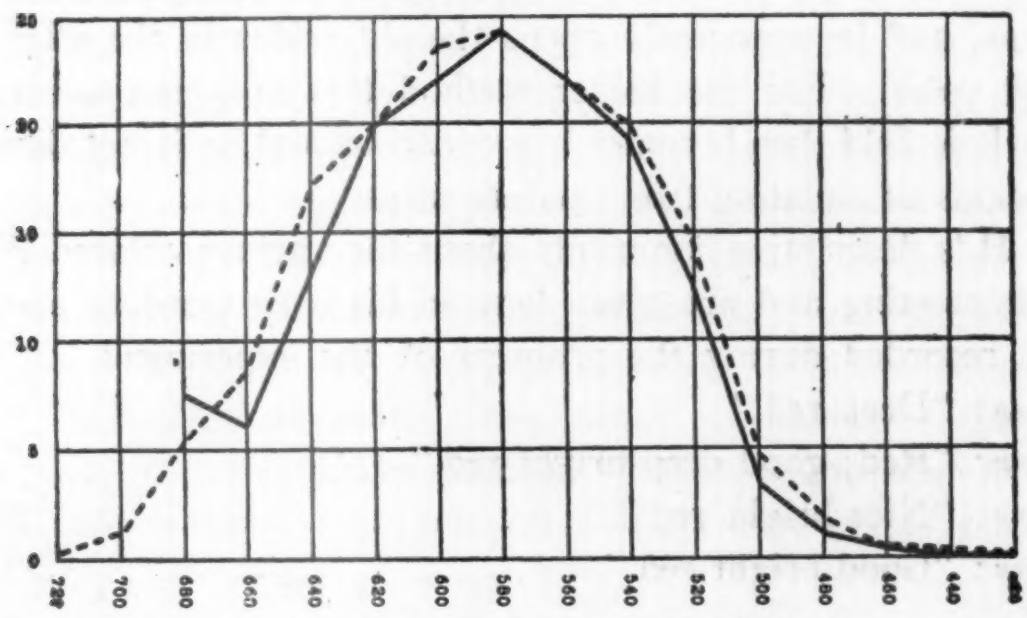

FIG. 4. Photopic luminosity curve for L. T. T. Right eye. - - - - - = direct comparison, ——— = flicker method.

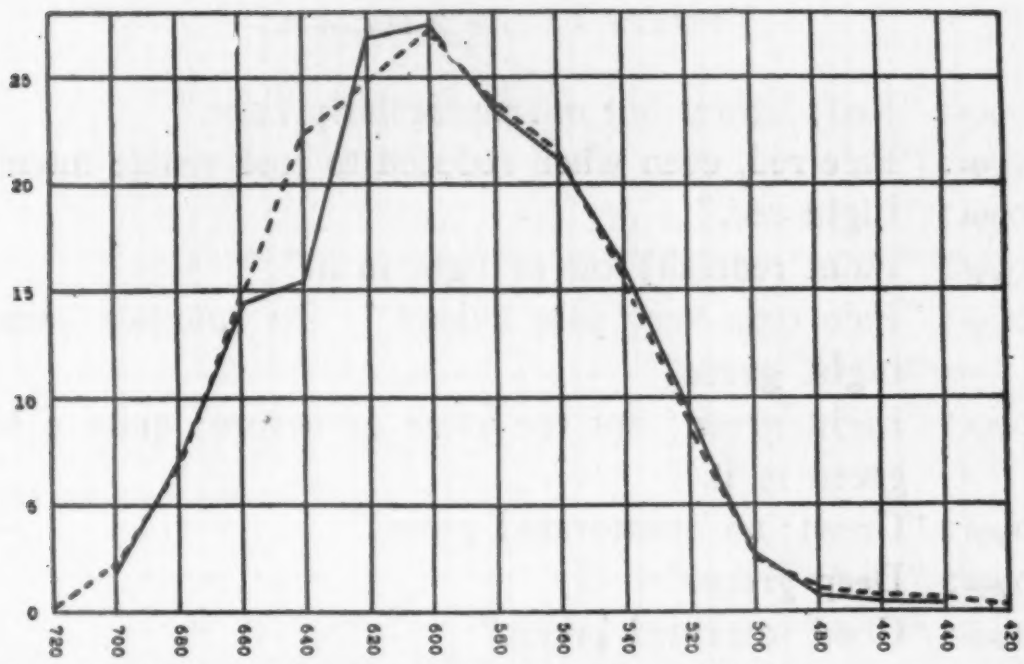

FIG. 5. Photopic luminosity curve for E. S. Right eye. - - - - - = direct comparison, ——— = flicker method.

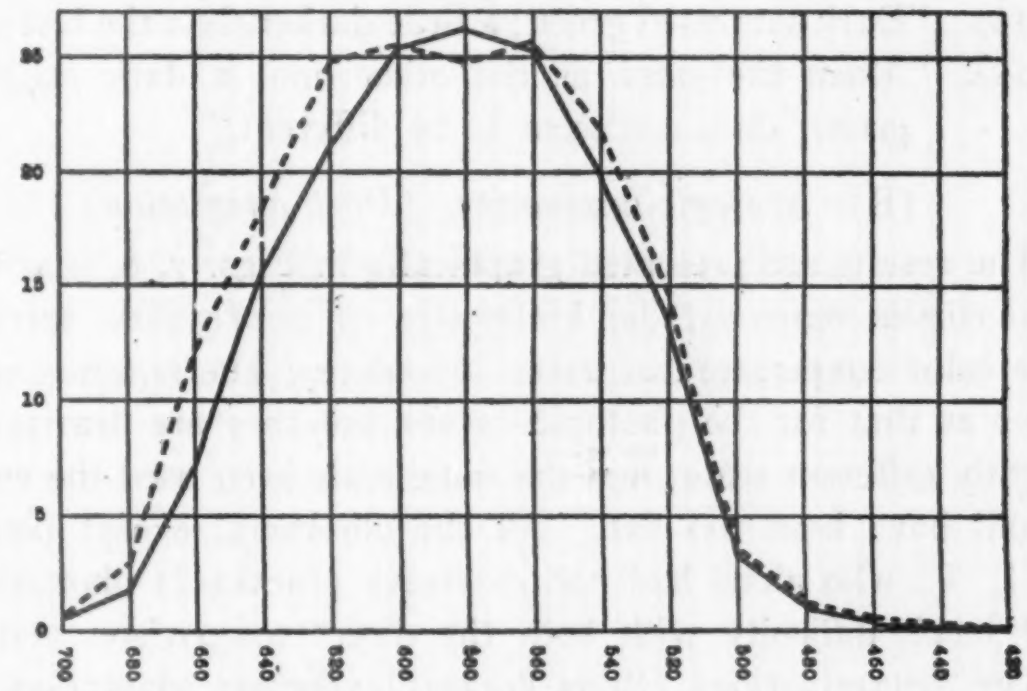

FIG. 6. Photopic luminosity curve for M. C. W. Right eye. - - - - - = direct comparison, ——— = flicker method.

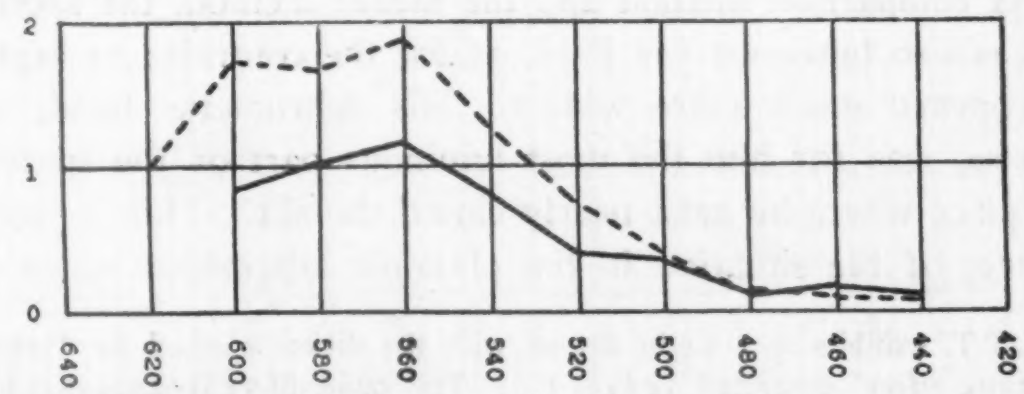

FIG. 7. Scotopic luminosity curve for J. H. Right eye. - - - - - = direct comparison, ——— = flicker method.

640 μμ: "Red; lighter but not particularly light."
620 μμ: "Pale red, even when reduced to just visible intensity."
600 μμ: "Light red."
580 μμ: "Faint reddish; lots of light in it."
560 μμ: "Pale red; very pale indeed." On another occasion, "Light green."
540 μμ: "Light green; not the same as before; quite a bit of green in it."
520 μμ: "Green; an unsaturated green."
500 μμ: "Deep green."
480 μμ: "Good saturated green."
460 μμ: "More saturated green than ever. A good deal more so than the last one."
440 μμ: "Dark saturated green; a little darker than the last one."
420 μμ: "Much the same as the other one; a dark saturated green; does not seem to be different."

(B). *Scotopic luminosity. (Dark adaptation)*

The results are presented graphically in Figs. 7, 8, 9 and 10. The curves represent the luminosity of a prismatic spectrum at a color temperature of 2100° Absolute. The notation is the same as that for the photopic curves, but they are drawn on a slightly different scale; had the same scale been used the curves would have been too flat. All the observers, except perhaps L. T. T., who alone had had extensive practice in photometry, had much difficulty with both the direct comparison and the flicker determinations. It is doubtful whether adaptation was complete enough to show the Purkinje effect.[6] With both the direct comparison method and the flicker method, the spectrum was not so luminous for J. H. as for the controls; he kept the slit opened much more widely. His achromatic band, near 560 μμ, was for him the most luminous part of the spectrum, the place where he most nearly closed the slit. This is just the reverse of the situation in the photopic adaptation curve with

[6] L.T.T. exhibits to a slight degree with the flicker method the "reversed Purkinje effect" described by Ives (4). The same effect is suggested in the case of J.H., and to a lesser degree with E.S.

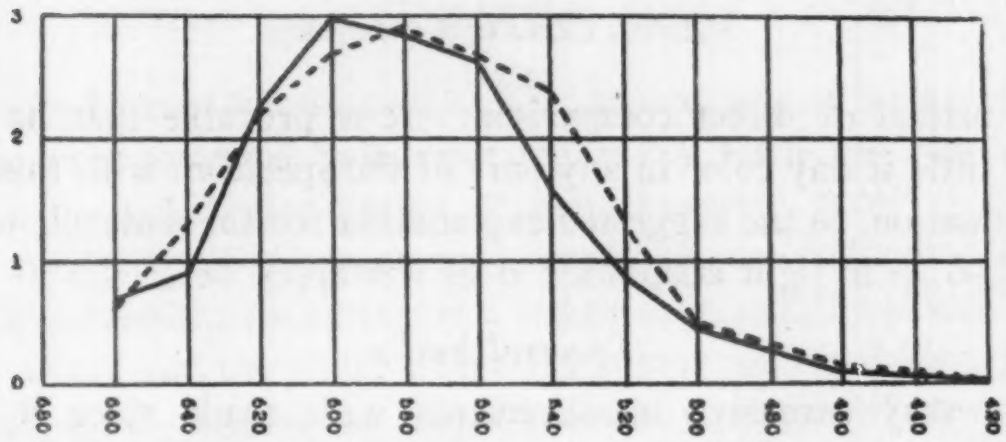

FIG. 8. Scotopic luminosity curve for L. T. T. Right eye. - - - - - = direct comparison, ——— = flicker method.

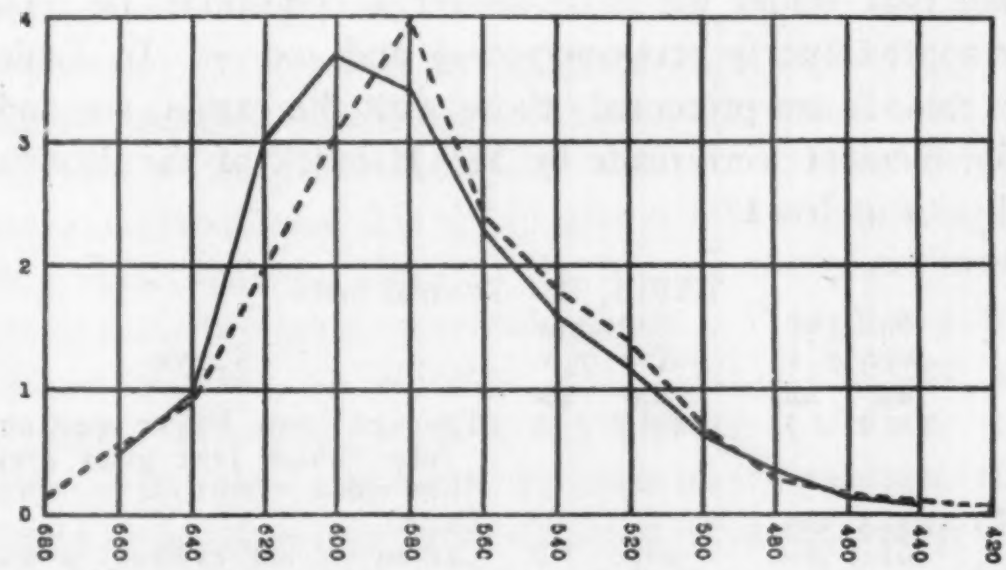

FIG. 9. Scotopic luminosity curve for E. S. Right eye. - - - - - = direct comparison, ——— = flicker method.

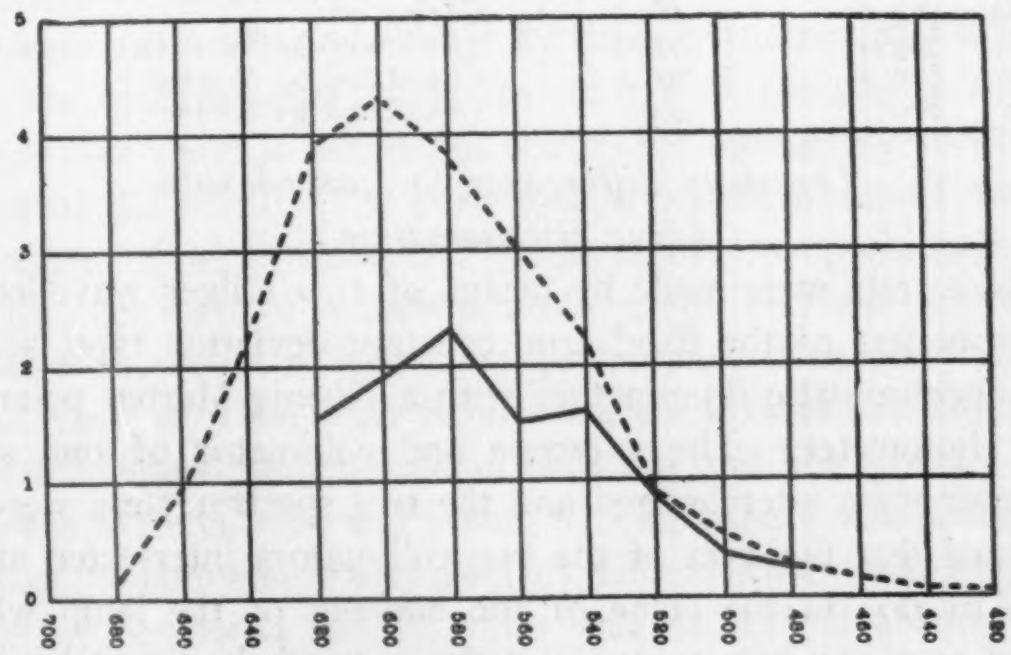

FIG. 10. Scotopic luminosity curve for M. C. W. Right eye. - - - - - = direct comparison, ——— = flicker method.

the method of direct comparison. It is probable that he saw very little if any color in any part of the spectrum with this low illumination, so the suggested explanation for the wide slit-width for 560 μμ in light adaptation does not apply here.

Spectral limits

No very extensive measurements were made since it was necessary to know only in a general way whether J. H.'s spectrum was shortened, especially at the long wave end. It may be said that, under the most favorable conditions, he reported color approximately between 750 μμ and 400 μμ. In Table III some records are presented; those with the carbon arc and the Nernst filament were made by Mr. Helmick of the department of physics at Iowa.

TABLE III—*Spectral limits*

	Red-end		Violet end		
	Ave.	m. v.	Ave.	m. v.	Source
Obs.	μμ	μμ	μμ	μμ	
J.H.	740.0	3.7	406.8	.9	Tungsten lamp, Hilger spectrometer, ruby Schott Jena glass over slit for red.
M.C.W.	767.5	4.3			Same.
J.H.	661.3	6.0	461.9	7.7	Carbon arc and ordinary prism.
R.H.S.	693.0	5.0	428.0	4.0	Same.
C.J.K.	701.0	8.0	411.0	4.0	Nernst filament.
P.J.S.	682.0	8.0	430.0	4.0	Carbon arc.
M.C.W.	712.0		404.0		Nernst filament.
M.C.W.	704.0		384.0		Carbon arc.
Norm	723.0		397.0		(As stated by Parsons.)
Norm	760.0		396.0		(As stated by Watson.)
Norm	800		400		(As stated by Köllner.)

Threshold differences for wave-length (*Hue discrimination*)

These tests were made by means of two Hilger wave-length spectrometers of the fixed arm, constant deviation type, set up as monochromatic illuminators with a Koenig-Marten polarization photometer. The telescope and collimator of one spectrometer were interchanged and the two spectrometers were so mounted that the axes of the two collimators intersected at an angle of 90° in the plane of the filament of the lamp which served as a common source, and the axes of the two telescopes similarly intersected a short distance in front of the slit of the

Koenig-Marten photometer. The source was a tungsten filament in a gas-filled bulb, operated at 0.5 watts per candle and having a color temperature of approximately 3000° Absolute. An image of the filament was formed on the collimator slit of each instrument by means of a simple lens suitably placed to fill the prism, and the width of the collimator slit was kept constant at 0.3 mm. The ocular slits were removed from the two spectrometers, and the photometer so placed that its slit served in their stead. The image of one collimator slit in the minimally deviated wave-length was formed directly on the upper half of the photometer slit; and, by reflection at a right angle prism, a similar image of the other collimator slit was formed on the lower half of the photometer slit. A piece of colorless ground-glass was placed over this slit to give a field of uniform brightness. After passing through the slit of the photometer, the light passed through a collimating lens, a Wollaston polarizing prism, a bi-prism, an objective lens, a Nicol analyzer, and an eye-piece. The field, as observed through a circular opening 1 mm. in diameter and 1 cm. in front of the eye, was circular, subtending a solid visual angle of about 3°, and bisected horizontally. The upper half was filled with homogeneous light from one illuminator, and the lower half by homogeneous light from the other illuminator. The wave-lengths selected at the photometer slit were determined by proper rotation of the prisms of the spectrometers. The two halves of the photometer field were kept equal in brightness during the settings by proper rotation of the Nicol. The experimenter set the prism of one spectrometer so as to illuminate the standard, or upper, half of the field with light of a known wave-length, and the observer adjusted the drum of the other prism table with one hand, and the Nicol with the other hand, until the variable, or lower, half-field matched the standard in both hue and brightness. Except in the extreme blue and violet, in which there was a slight disproportionality in the amount of diffuse white light superposed on the compared spectra, there was no difficulty in setting so that the entire field appeared uniform. The field was of a high

brightness between the extremes of the spectrum. It was not feasible, owing to the stringency of time, to keep the standard half-field constant in brightness for all wave-lengths. The results given by the several observers were obtained under identical visual conditions and the results are strictly intercomparable. The observing room was in semi-darkness.

The procedure was as follows: Beginning at the extreme red, 700 $\mu\mu$, the two spectrometers were set so as to illuminate both halves of the field in the same wave-lengths. The observer was required to keep the two halves equal in brightness, and to turn the prism of the instrument illuminating the variable half-field until it appeared just noticeably different in hue from the standard. The reading was recorded by the experimenter; the difference was increased, and the observer required to turn the prism until the difference in hue just disappeared. Five paired settings were taken in this way, and the wave-length illuminating the standard half-field was changed to the mean of the ten readings to serve for the next standard. In this manner, the whole spectrum was worked through in short sittings, with frequent rest periods, for three observers, J. H., and two controls, E. S. (who also served in the luminosity experiments) and the writer. In the case of J. H., a second set of readings was taken, working from the violet, 410 $\mu\mu$, to the red.

The results are presented graphically in Fig. 11, the long, folded chart. Along the base lines the 10 $\mu\mu$ divisions are indicated. Normal spectral limits are marked by the heavy lines (Fraunhofer, A and K) at either end. Since the Fraunhofer lines afford fixed points for reference, they are drawn in, with their wave-lengths and characteristic hues indicated. The small crosses give the initial starting points for each observer, and the arrows the direction followed. The short lines, perpendicular to and above the base, represent, in position, the discrimination intervals, each one marking the hue that is just perceptibly different from the preceding one under the conditions imposed. The height of these lines above the base shows the mean variation characterizing the average of the 'difference' and 'equality'

settings; a height of 1.5 mm. being approximately equal to 1 $\mu\mu$ of variation as the chart appears after reduction. The initials at the left indicate the three observers.

Inspection of the curve reveals at once the vast difference between J. H. and the two control observers; for him hue discrimination practically does not exist. In his first record there are twelve recorded changes and in the second there are only six.[7] M. C. W. has under exactly similar conditions eighty-eight hue changes, and E. S. seventy-two. For J. H. there is no change from 700 $\mu\mu$ to 583 $\mu\mu$; "red" therefore is constant and unchanging in hue until well into the yellow. The few changes that then occur mass themselves for the most part within the next forty or fifty $\mu\mu$. At about 566 $\mu\mu$ a very narrow colorless or achromatic band appears, the limits for which have been very carefully determined. Passing on through the green which he calls "green" and into the blue which he also calls "green", changes are infrequent, until the violet, remarked as "green with some red in it," begins to appear, and continuing, color is finally lost, and "just light" remains. Could his spectrum be bent around into a circle, it would then show two achromatic bands; the one near 566 $\mu\mu$ already mentioned, and one between the violet and the red. The second record for J. H. is not qualitatively different from the first.

Hue changes for M. C. W. mass in two regions, the orange-yellow and blue-green. E. S. shows evident practice effects (supported by his verbal comments during the experiment) and his differences grow smaller rather gradually, but show a tendency to mass in the blue-green region also.

Neither M. C. W. nor E. S. attain the number of hue differences calculated by Koenig from Uhthoff's results, (5), namely 165. With practically the same apparatus that was used in this experiment, Dr. Johnson informs the writer that Nutting, in some unpublished experiments, obtained about 200 hue differences. Both J. H. and the controls however were quite without practice in this particular experiment.

[7] This second record was made very hastily, just before J.H. had hurriedly to leave for the train, and it may have been influenced by that fact.

In connection with this experiment, J. H. made an observation which is interesting. It was noticed that he looked longer than usual at the color field with a certain setting of the spectrometers, (upper half of field 551.7 $\mu\mu$, lower half 554.0 $\mu\mu$ and somewhat darker). Asked what the difficulty was he replied: "It is beautiful and reminds me of an ocean sunset view at the equator. There is the sky all glowing red and the green ocean below. I have seen it that way many times." To the experimenter, who took a look at this seascape, there was a greenish yellow "sky" and a somewhat darker dirty yellow "ocean", and no straining of the imagination could make out of this a brilliant marine sunset comparable to those she had actually seen.

Matches between mixtures of two homogeneous lights and a single homogeneous light

The dichromat is usually able to match any spectral color by mixtures in suitable proportions of light from the extremes of the spectrum. A Hilger wave-length spectrometer of the fixed arm, constant deviation type, was set up as an illuminator so that one of the opal glass receiving surfaces of a photometer was illuminated by the wave-length suffering minimum deviation by the prism of the spectrometer and selected at its ocular slit. A high intensity monochromatic illuminator was used to fill the other half of the photometric field. This instrument was equipped with a train of two 60° dense flint glass prisms which were set for minimum deviation of $\lambda = 505$ $\mu\mu$. The spectrum was formed on an acid-blackened brass plate containing three slit openings, the width and positions of which were independently variable. One of the slits was closed and the others placed so as to transmit a band at 672.6 $\mu\mu$ and another at 461.4 $\mu\mu$.[8] A projection lens was used to form an image in these two bands of the face of the first prism and the second receiving surface of the photometer.

The source used for both illuminators, were tungsten filaments in argon-filled bulbs, operated under steady current con-

[8] The nearest fixed points to these hues are the red lithium at 670$\mu\mu$, and the blue helium at 471.3$\mu\mu$.

ditions and giving a color temperature of approximately 3000° Absolute.

The field as observed was circular and subtended a solid visual angle of about 4°. Its mean brightness was high enough for comfortable reading, but was not measured. It was observed in dark surroundings through a circular opening 1 mm. in diameter and 1 cm. from the eye. The pattern was two half-circles each with a trapezoid. In all cases a match was obtained in which the pattern disappeared against the background.

The data secured in the mixing of spectral lights are not so extensive as had been planned, nor is the method of procedure so precise, for an unavoidable three-day absence from Cleveland made extensive work with the color mixing apparatus impossible. The original plan had been to proceed by taking definite proportions of the red and blue lights and find the monochromatic equivalents, working from 100% red + 0% blue to 0% red + 100% blue. What was actually done, however, was to take 10 $\mu\mu$ stops of monochromatic light along the spectrum, and determine empirically the proportions of mixed red and blue beams which would match for J. H. each monochromatic beam. An exact match was insisted upon, however; the pattern in the field (mixed light) was made to match the background (monochromatic) so that the former disappeared. The slit-width of the Hilger spectrometer illuminating the background was kept constant, so the luminosity curve is incorporated in the records. This means that, for the intense colors, wider, therefore more intense, red and blue beams had to be admitted.

An attempt was made to get a rough estimate of the relative stimulating values of the red and blue lights in terms of the same constant. The Hilger spectrometer was set by J. H. at the place where the background illuminated by it showed the least color, namely at 564.5 $\mu\mu$, slit-width 0.5 mm. A direct comparison luminosity match for this was found with red, 667.8 $\mu\mu$, to require a slit-width of 0.5 mm. The same for the blue, 460.3 $\mu\mu$, was 4.4 mm. Translating slit-width into wave-lengths, $\mu\mu$, this means that a beam of red light extending from 673.8 $\mu\mu$

to 667.8 μμ is approximately equal in stimulating value to a blue beam extending from 476.4 μμ to 460.3 μμ, since both, with unchanging light source, are apparently equal to the same unchanging monochromatic beam. Since hue discrimination practically does not exist for J. H., it may be assumed that the narrow beams here admitted were of one hue throughout.

TABLE IV

	Red			+ Blue		
Stand.	R. Jaw	L. Jaw	Range	R. Jaw	L. Jaw	Range
μμ	μμ	μμ	μμ	μμ	μμ	μμ
460	None			461.4	463.0	1.6
470	None			461.4	463.2	1.8
480	None			461.4	463.3	1.9
490	None			461.4	467.0	5.6
500	None			461.4	469.5	8.1
510	None			461.4	473.0	11.6
520	672.6	672.8	.2	461.4	473.5	12.1
530	672.6	672.9	.3	461.4	473.5	12.1
540	672.6	674.2	1.6	461.4	473.2	11.8
550	672.6	674.5	1.9	461.4	472.5	11.1
560	672.6	675.0	2.4	461.4	472.3	10.9
570	672.6	687.2	14.6	461.4	472.2	10.8
580	672.6	714.0?	41.4		None	
590	672.6	731.0?	58.4		None	
600	672.6	730.0?	57.4		None	
610	672.6	724.0?	51.4		None	
620	672.6	710.0?	37.4		None	
630	672.6	703.0	30.4		None	
640	672.6	702.0	29.4		None	
650	672.6	698.0	25.4		None	
660	672.6	685.0	12.4		None	

In Table IV the best one of three complete sets of records is given. The table is practically self explanatory. 'Range' refers to the width of the monochromatic beams in μμ admitted between the jaws of the slits allowing the selected beams to pass. The wave-lengths at each jaw of each slit are given merely for the sake of completeness. The interrogation marks indicate readings carried beyond the point of accurate calibration.

It is very desirable to plot the results of this experiment in a curve, but since the stimulating values of red and blue are so unlike, that cannot be done directly, nor do the records give for J. H. all the values necessary for making the calculations, and further it is probable that, although every effort was made to keep the physical conditions in the apparatus constant, many variables entered. An attempt was made to plot such a curve

based upon certain calculated results; it is not presented on account of probable unreliability.

Von Kries (6) has emphasized the fact that a mixture of red (670.8 $\mu\mu$) and yellow-green (550 $\mu\mu$) will match for protanopes, deuteranopes and normal 'trichromats' a homogeneous yellow, although the three classes do not agree among themselves as to the proportions of red and green in the mixtures. Von Kries, however, does not keep the hue of his yellow constant, since in the table he gives, the wave-length varied from 569 $\mu\mu$ to 639 $\mu\mu$. Although this test is not an especially critical one, the attempt was made with J. H. to match a homogeneous yellow (587.6 $\mu\mu$ yellow helium) with a mixture of red (667.8 $\mu\mu$, third red helium) and green (501.6 $\mu\mu$, second green helium). Intensity of the yellow was varied by changing the slit-width. The mixed light illuminated the pattern of the field, the monochromatic light the background. A direct comparison luminosity equivalent of yellow 587.6 $\mu\mu$, slit-width 0.5, was found in the green, 501.6 $\mu\mu$, slit-width 1.4 mm.; and in the red, 667.8 $\mu\mu$, slit-width, 3.07 mm.

No attempt has been made to reduce the results to equal stimulating value. In Table V, the amount of each of the two colors in the mixture as empirically determined is stated in terms of the width of the beams in $\mu\mu$'s passed through the slits. The right jaws of the slits were fixed respectively at 667.8 $\mu\mu$ and 501.6 $\mu\mu$.

TABLE V

Yellow 587.6$\mu\mu$ Slt-wdth. mm.	Green 501.6$\mu\mu$ Range $\mu\mu$	+ Red 667.8$\mu\mu$ Range $\mu\mu$
.8	1.5	71.2
.7	1.9	32.2
.6	1.6	17.2
.5	1.4	35.2
.4	1.6	37.2
.3	1.8	20.2
.2	1.5	10.2
.1	.7	4.2

The table shows that J. H. accepts relatively only a small amount of green in the mixture, and that this amount remains

a fairly constant quantity as the luminosity of the yellow is successively reduced. The luminosity adjustment is made in the red. J. H. would rather have had no green in the mixture; he detected its presence and requested its withdrawal when the amount was increased. The yellow at 587.6 $\mu\mu$ is for him distinctly within the red, very near to the point which he selects as the "prettiest and brightest red in the spectrum", namely at about 600 $\mu\mu$, and might easily have been matched by different amounts of red alone. To the normal eye 587.6 $\mu\mu$ is practically an orange yellow.

The achromatic band

It was early discovered there was an apparently colorless band at about the place in the spectrum where the yellow begins faintly to appear for the normal eye. Its position was first rather definitely known when J. H. set an Hilger spectrometer to 564.5 $\mu\mu$ as showing the "least color". Later a slit with bilaterally movable jaws was put in the path of the spectrum at the place of greatest achromaticity, and the slit opened until the edges just marked the limits of this region. The average of five trials, fixed the limits at 568.2 $\mu\mu$ and 564.3 $\mu\mu$, indicating a band about 4 $\mu\mu$ wide, with the center at 566.2 $\mu\mu$.

Later on, a still more exact determination was made by adapting the hue discrimination apparatus. One-half of the field of the polarization photometer was illuminated with light from an Hilger instrument, the other half by a small tungsten automobile headlamp operated at 6 volts. Brightness equality was easily secured by rotating the Nicol prism. The Hilger drum which rotated its prism table was set first by the observer at a point where no color appeared when compared with the approximately white standard half-field, then the drum was turned toward the red until red just began to appear and the reading was taken by the experimenter. The drum was then rotated in the reverse direction until color ("green") began to come in from the opposite side. Five settings for each border limit were secured, which, when averaged, gave 568.38 $\mu\mu$ and 566.52 $\mu\mu$ as the limits. The range is 1.86 $\mu\mu$, and the center is located

at 567.45 μμ. This is in very close agreement with the other determinations of the colorless band, and not far different from J. H.'s unaided setting of the Hilger at 564.5 μμ as the place of least color. From the band to the long end, the spectrum was called "red", and in the other direction it was "green" throughout, with no extensive shortening at either end. As has been noted previously however, were the spectrum bent in a circle, there would be another achromatic region between the violet and red.

Miscellaneous

Time did not permit definite experimental study of after-images and simultaneous contrast. Incidental observations were made from time to time and indicate that both these effects occur. In the case of contrast, when one half of the field was white, the other yellow, ("red"), some "green" was detected in the white.

White was repeatedly called white, and aside from obvious induction effects was never confused with a color.

The best, purest and brightest "red" he located at about 600 μμ; the best "green" at about 450 μμ, (M. C. W. best blue at 440 μμ). His "red" is thus really in the orange-yellow and his "green" in the blue.

What effect macular pigmentation has in this case is not known. His whole retina, including his macula, is rather densely pigmented, but it is not exceptional in this respect. Parsons says that the effect of macular pigmentation is to displace the achromatic band toward the long end, in this case, into the yellow. Color is probably not intense for J. H. in this region of the spectrum and macular absorption may be rather effective.

At various times, incidental tests were made to determine whether there was a marked difference between the two eyes. No noteworthy difference was found.

This study was begun, and for the most part pursued, without any conviction as to the type of color defect represented by J. H. It was sufficient to know that the case was unusual to stimulate

the investigation. The study is not complete, but it is not certain that J. H. will be available for further work. Color theories are at the present time undergoing modification and new data are needed. The writer, then, presents these as the facts in hand, and submits the description without at this time attempting a diagnosis.

REFERENCES

1. Hayes, S. P. The color sensations of the partially colorblind; a criticism of current teaching. *Amer. J. of Psychol.*, XXII, 381.
2. Lovibond, J. W. *Measurement of light and color sensations.* London. Geo. Gill and Sons, 1893.
3. Ives, H. E. Studies in the photometry of lights of different colors. *Phil. Mag.*, XXIV, 1912, 149 ff.
4. Ives, H. E. Ibid, 170.
5. Parsons, J. H. *An introduction to the study of color vision.* G .P. Putnam's Sons, 1915, 32.
6. Von Kries. Über Farbensysteme. *Zsch. f. Psychol. u. Physiol. d. Sinnesorg.*, XIII, 277.

BINAURAL BEATS

BY

G. W. STEWART

CONTENTS

Introduction.
Earlier results.
Observations with F. C. Bruene.
Observations upon the secondary maxima.
A Theory of binaural beats.
  The primary maxima.
  The secondary maxima.
  The localization.
Anatomical considerations.
Restatement of theory and agreement with experiment.

Introduction

The subject of binaural beats has been one of much interest to the psychologist,[1] not only because the phenomena observed are difficult to explain with entire satisfaction, but also because the explanations involve what are evidently very important questions. The writer became interested in the physical aspects of binaural hearing through his theoretical and experimental studies[2] of the sound intensity and phase relations in the neighborhood of a rigid sphere. His experimental work upon the subject of binaural beats will be presented in appropriate detail in the Physical Review for 1917. The present article will give a very brief account of the results, both experimental and theoretical, with the expectation that those especially interested in this field will refer freely to the detailed reports in the Physical Review. Brevity will necessitate the preparation of this article for the most part as an abstract, and hence omissions of refer-

[1] An excellent review of the experiments with bilateral excitation is given by Rostosky, Beiträge zur Psychologie and Philosophie, I, 1905, pp. 173-273.

[2] G. W. Stewart, Physical Review, (1911), and N. S., (1913), (1914) and (1916).

ences to other investigations and of complete discussions of disputed points, will occur frequently.

Earlier Results

The chief phenomena observed when two tones differing in frequency are led one to each ear are as follows:[3]

(a) Beats occur whether the tones are presented to one ear or one to each ear. This was discovered by Dove,[4] observed subsequently by various experimenters and discovered independently by S. P. Thompson.[5] This phenomenon is observed by all, even by the partially deaf.

(b) With bilateral presentation of the tones, the minimum intensity observed in each beat does not (even with carefully adjusted equal intensities of the two sources) approximate zero. This is mentioned by S. P. Thompson[6] but doubtless was observed very much earlier by others.

(c) When the tones at the ears are alike in intensity and phase, there is a localization in the median plane, the position in this plane varying with the observer and presumably with the apparatus arrangement. Purkyné[7] was the first to record this localization, which has been shown through numerous experiments of other observers, to be a usual accompaniment of binaural perception.

(d) With opposition in phase and equality in intensity, the sound is localized according to S. P. Thompson[8] at the back of the head. To several observers, with apparatus of a form different than that of Thompson's the localization is in the median plane. Doubtless the position in the median plane is influenced by the method of observation.

(e) The sound is localized on the side of the fork leading

[3] Unless otherwise stated, the difference of freqnency is assumed to be small.

[4] Loc. cit.

[5] S. P. Thompson, Phil. Mag. 5, IV, 1877, p. 274.

[6] Loc. cit.

[7] Purkyné Prager Vierteljahrsschrift, Bd. 67, p. 91, 1860.

[8] S. P. Thompson, Phil. Mag. 5, IV, 1877, p. 274, and Phil, Mag. 5, VI, 1878 p. 383.

in phase when the phase difference is approximately 90°. Thus with phase difference the sound wanders from the median plane at 0° to one side at 90°, to the median plane at 180°, to the other side at 270° and again to the median plane at 360°. Perhaps S. P. Thompson[9] first observed the wandering from ear to ear but the periodic displacement of the localization was first most clearly described by Paul Rostosky.[10] Later workers, who were not aware of Rostosky's results, observed this lateral displacement. Lord Rayleigh,[11] experimenting with forks, recognized the lateral effect and its relation to phase difference, and later[12] confirmed these observations with experiments conducted by the use of telephones. L. T. More and H. S. Fry[13] made preliminary experiments which were followed by those of L. T. More.[14] His results showed unmistakably the lateral displacement. More used an apparatus similar to that earlier adopted by C. S. Myers and H. A. Wilson.[15] A single source, branched pipes to the ears and an arrangement whereby the lengths of the pipes could be altered in opposite senses, were used. The experiments of Myers and Wilson also gave full confirmation of the lateral effects. Bowlker's[16] experiments do not seem to furnish reliable evidence. Unfortunately, the workers just mentioned seem not to have been aware of the splendid experiments of Rostosky published in 1902.[17]

(f) There is a limit to the frequency of the tone with which the effect just described (e), can be obtained. L. T. More[18] states that a fork of 512 d. v. was near to his limit of accuracy in the judgment by phase differences, that with 1024 d. v. his judgment became untrustworthy and that with 3000 d. v. he

[9] S. P. Thompson, Phil. Mag. 5, XII, 1881, p. 351.
[10] P. Rostosky, Philosophische Studien, 19, 1902, p. 557.
[11] Lord Rayleigh, Phil. Mag. XIII, 1907, p. 214.
[12] Lord Rayleigh, Phil. Mag. XIII, 1907, pp. 316-319.
[13] L. T. More and H. S. Fry, Phil. Mag. XIII, 1907, p. 452.
[14] L. T. More, Phil. Mag. XVIII, 1909, p. 308.
[15] C. S. Myers and H. A. Wilson, Proc. Roy. Soc. LXXX, 1908, p. 260.
[16] T. J. Bowlker, Phil. Mag. (6) XV, 1908, p. 318.
[17] Loc. cit.
[18] Loc. cit.

had no sense whatever of direction. Lord Rayleigh[19] gives the limiting frequency of his right and left sensations at 768 d. v. C. S. Myers and H. A. Wilson[20] state in general terms that with very high frequencies the lateral effects cannot be obtained.

(g) In addition to the maximum at 0° phase difference, there are also two additional maxima occurring one before and the other after opposition in phase, but distinctly perceptible only when the beats are less frequent than one in two to five seconds. These additional maxima were first studied by Paul Rostosky[21] in 1902, but in the succeeding literature no reference is made to his important contribution. Thus the subsequent discovery of these secondary maxima by the writer was an independent one.

Observations with Mr. F. C. Bruene

The tones were obtained with two tuning forks, the sounds being led to the ears by binaurals inserted[22] in the opening of the external meatus. Phase relations were determined by the Lissajou's figures made by a beam of light reflected from two mirrors, one attached to each fork.

The following results were obtained with the assistance of Mr. F. C. Bruene, a graduate student in the Department of Philosophy and Psychology.

(1) The beats (i. e., maximum at 0° and minimum at 180° phase differences) with the tones of two forks, one applied to each ear, were heard distinctly. Indeed, of twenty-three inexperienced observers listening for three periods of five minutes each, seventeen noticed the beats distinctly enough to report them. All experienced observers can hear them and record them with very satisfactory accuracy.

(2) There existed a distinct wandering of the localization; in front at 0° difference of phase, on the side of the fork with the higher frequency leading in phase from 0° to 180°, and on

[19] Lord Rayleigh, Phil. Mag. XIII, 1907, p. 214.

[20] Loc. cit.

[21] Loc. cit.

[22] See section VII in which there is a discussion of the use of binaurals closing the external meatus.

the side of the slower fork with the higher frequency leading in phase from 180° to 360°, the changes of position being continuous, from front to rear, (or within the head), and from rear to front. This wandering when first recognized might be described as occurring in a circular path. A more accurate description occurs in what follows.

(3) This rotation of the localization just described was quickly and distinctly observed by some hearers; on the other hand all observers did not recognize it, even with an hour or more practice. Eleven of the twenty-three uninstructed and inexperienced observers reported their observation of this "rotation" after fifteen minutes' trial (three trials of five minutes each). Six of these eleven noticed the effect within the first five minutes, and three in the second.

(4) All of the eleven just mentioned, with the possible exception of one or two, noticed that the localization was much more distinct in the half cycle in which the lead of the faster fork varies from —90° to +90°. With experienced observers, the effect is very noticeable.

(5) Extended experiments with four experienced observers showed that the most accurate judgment as to phase difference, made by the localization, occurred when that difference is 0°.

(6) With the experienced observers the change in the localization at 0° phase difference by the production of a very marked inequality in intensities, is slight. The rotation of the localization was modified by a slight circular shift in the direction of the side of the greater excitation. The effect upon the localization at the sides became very marked, that on the side of the greater intensity becoming more prominent and that on the other almost vanishing.

(7) With observers of experience, but who had not yet recognized the secondary maxima,[23] the maximum intensity seemed to occur at 0° phase difference when the beat-period was less than perhaps two seconds. With a greater and gradually increasing period there arose an uncertainty until, with a period

[23] See pp. 35, 37 and 41 of this article and (9) below.

of perhaps five seconds, the judgment of maximum intensity shifted to a phase difference noticeably less than 180°, or approximately to the first occurring secondary maximum.

(8) The rotation of localization with experienced observers occurred as follows: At 0° phase difference the sound was localized as in front and distant, i. e., external to the head. The localization then described in a horizontal plane a path which appeared to be somewhat circular, but in which the apparent distance contracted until the path entered the ear leading in phase. Then at about 180° phase difference the localization passed quickly through the head, into the other ear and from thence around to the front along a path symmetrical to the one just described.

(9) When an experienced observer gave his attention to apparent intensity rather than to localization, there appeared three fairly distinct maxima. One occurred at 0° difference of phase, one at 180°— δ and one at 180°+ δ, δ being less than 45°. The earlier of the additional maxima at 180°— δ phase difference coincided with the localization in the ear near the higher frequency, and the later one with the localization in the other ear. With a 128 d. v. fork the secondary maxima seemed clearly discernible only if the beat-period exceeded a value of approximately two seconds. The value given is only approximate as the variation in the perception of the secondary maxima is continuous with the changing of the beat-frequency. The significant fact is that the 0° maximum was always present, whereas the secondary ones required a beat-period exceeding at least a second in order to become clearly evident.

(10) The localization in the ears as described seemed to depend upon the perception of the secondary maxima. If they were absent, as at high beat-frequency, the localization seemed confined to the region near the front. In other words the localization in approximately one-half the cycle seemed to be due to the secondary maxima or the causes producing them.

Observations upon the Secondary Maxima

Again the binaurals inserted[24] in the ears were used. The observations were made with the assistance of Dr. Harold Stiles of Iowa State College. As already stated, the secondary maxima are those occurring just before and after the 180° difference in phase, or at phase differences of 180°—δ and 180°+δ. The results were obtained by extended observations upon pairs of forks of the following frequencies: 42, 63, 128, 256, 361, and 469 d. v. The length of the beat-period and time between the secondary maxima were measured. From these values $\frac{\delta}{\pi}$ was computed. The results are shown in Tables I, II and III. The number given in parentheses after the values of 2δ in Table I and after the values of $\frac{\delta}{\pi}$ in Tables II and III, is the actual number of observations involved in the value given of 2δ or $\frac{\delta}{\pi}$. The conclusions are as follows:

(11) The difficulty of determining the exact moment of the occurrence of a maximum renders the measurement of δ inaccurate. Observations were, however, made at various times of day and the average results do seem to warrant the two conclusions that follow.

(12) For a given pair of forks δ is independent of the frequency of the beats when this frequency is varied from one to three times. This is shown especially in Table I.

(13) δ does vary with the fork frequency, and this variation is not linear, varying less rapidly than the frequency. This is shown by the averaged values taken from all the tables.

A Theory of Binaural Beats

The phenomena already mentioned are complex and hence we will divide the theoretical discussion into three parts, namely, the primary maxima, the secondary maxima and the localization.

[24] See footnote 22.

TABLE I

Fork of 128 d. v.

Date	Period of beats	2δ in seconds	$\frac{\delta}{\pi}$	$\frac{\delta}{\pi}$
Aug. 3	27.8		.20	
	12.3	2.60(7)		.21
	12.2	3.50(7)		.29
	13.1	2.75(8)		.21
Aug. 10	35.7	7.9 (4)	.22	
	35.5	7.2 (6)	.20	
	36.0	10.3 (9)	.29	
	13.7	2.95(12)		.22
	14.0	3.6 (11)		.26
Aug. 11	35.2	7.9 (4)	.22	
	13.3	2.68(10)		.20
	13.5	3.1 (6)		.23
	Mean		.23	.23

TABLE II

Forks of 42, 63, and 128 d. v.

Fork of 42 d. v. Date	Beat Period	$\frac{\delta}{\pi}$
Aug. 11	20.4	.19(10)
	20.2	.20(8)
Aug. 12	22.6	.18(10)
	23.7	.16(8)
	22.6	.18(9)
	24.5	.15(11)
	21.6	.17(13)
	21.5	.18(10)
	Mean	.17
Fork of 63 d. v.		
Aug. 3	10.1	.20(11)
	10.7	.20(13)
	10.8	.18(11)
Aug. 5	10.2	.18(4)
	19.5	.16(6)
	19.0	.14(4)
	Weighted Mean	.18

Fork of 128 d. v. Date	Beat Period	$\frac{\delta}{\pi}$
June 27	23.2	.25(2)
	10.1	.29(8)
June 29	23.9	.24(3)
	31.0	.22(3)
	16.0	.27(4)
July 2	35.8	.22(4)
July 3	27.8	.20(6)
	12.3	.21(7)
	12.2	.29(7)
	13.1	.21(8)
July 10	36.0	.29(9)
	35.5	.20(6)
	35.7	.22(9)
	13.7	.22(12)
	13.7	.25(10)
	14.0	.26(11)
	13.6	.20(7)
July 11	13.3	.20(10)
	13.5	.23(9)
	35.2	.22(4)
	Mean	.23

The Primary Maxima.—If we assume that there is sound conduction across the head by paths not entirely aerial, and that the organ of hearing detects actual intensity, or an effect which depends upon the square of the amplitude of vibration, and if we further assume that there is a blending which results

TABLE III

Forks of 256, 361, and 469 d. v.

Fork of 256 d. v. Date	Beat Period	$\frac{\delta}{\pi}$
July 13	13.2	.26(9)
" "	12.2	.29(9)
" 14	43.8	.26(4)
" "	27.4	.27(8)
" "	23.4	.21(8)
" "	12.0	.28(11)
" "	10.5	.26(9)
" "	9.25	.27(9)
	Mean	.26

Fork of 361 d. v. Date	Beat Period	$\frac{\delta}{\pi}$
June 26	15.9	.28(9)
" "	8.2	.29(10)
July 20	17.6	.28(6)
" "	22.3	.26(7)
" "	17.1	.25(8)
" 27	10.7	.27(9)
" "	10.2	.24(15)
" "	11.7	.30(5)
" "	9.1	.28(9)
" "	12.5	.22(7)
Aug. 1	12.7	.30(6)
	Mean	.27

Fork of 469 d. v.		
Aug. 5	4.2	.32(1)
" "	4.0	.37(2)
" "	6.6	.27(4)
" "	5.6	.35(2)
" "	5.0	.38(2)
" "	5.0	.26(3)
" "	7.0	.27(3)
	Mean	.32

in the effects at the two ears appearing as one, we obtain the result that the total sensation must vary with the following expression:

$$2(\alpha^2+\beta^2)+4\alpha\beta\cos\epsilon\cos\lambda$$

where α and β are constants, ϵ is the phase by which the faster fork leads the slower, and λ is the difference between the phases introduced in the two paths, one from the source to the nearer ear and the other from the source to the farther ear. Thus ϵ is the only variable in the expression, and hence the total sensation must pass through a maximum and a minimum (not zero), the former corresponding to $\epsilon = 0°$ and the latter to $\epsilon = 180°$. According to the experimental results just presented, this theory agrees with the fact that the maximum and minimum do occur at the phase difference specified, the fact that the minimum is not zero; and the fact that phase differences for maxima and minima are independent of either fork-frequency or beat-frequency.

The Secondary Maxima.—The only theory proposed for the explanation of the secondary maxima has naturally been that of the only previous observer, P. Rostosky.[25] He contended

[25] P. Rostosky, Philosophie Studien, 1902, 19, p. 557.

that they were caused by the variation in intensity-ratio, but, unfortunately, in the interpretation of his equations, he made an error[26] which when corrected makes his theory contradict experimental facts.

In order to explain not only the secondary maxima, but various other related phenomena, we are compelled to make several assumptions in addition to the ones just stated in the preceding discussion. An inspection of the experimental results will convince that the secondary maxima must have an origin quite independent of the primary maxima. The assumption will therefore be that the organ of hearing is, from a physical point of view, duplex. A pair of "A" instruments is responsible for the primary maxima and a pair of "B" instruments is responsible for the secondary maxima. We accept, then, the existence of the B instruments and in addition make the following assumptions: that the blending with the B is not as perfect as with A instruments, that with the former there is introduced an additional change of phase of 180° in the wave transmitted across the skull and that the B instrument is excited by the skull rather than by the air-membrana tympani route. We find by a theoretical investigation that the intensity in the ear nearer to the faster fork reaches a maximum at $180° - \lambda$ and in the further ear at $180° + \lambda$, where λ is as previously defined. Thus we have explained the existence of the secondary maxima as the recognition of the individual intensities at each B instrument, and λ is identified as the δ observed experimentally. Thus the result stated under (g) is explained. If the B instruments are excited by the skull, which is comparatively massive, a certain length of time would be required to set up any vibratory motion in these B instruments. This is in accord with the experimental fact, as expressed in (9), that the secondary maxima are not in evidence unless the primary beat-period exceeds one or two seconds. The fact that δ is independent of beat-frequency as stated in (12) above, is due to the fact that the retardation λ (in the theory) depends only upon the frequency of vibration and the

[26] See article on "Theory of Binaural Beats," by G. W. Stewart, in the Physical Review for 1917.

conducting paths and not upon the rate at which the phase differences of the sources are changing, or upon the beat-frequency.

The fact that δ is dependent upon the fork-frequency, as stated in (13) above, is in agreement with the theory, for the time retardation along any path is constant and the phase retardation is strictly proportional to the frequency. Furthermore it can be shown[27] that when there are numerous paths between two points, the resulting phase retardation cannot in general be a linear function of the frequency. Thus our theory is in harmony, qualitatively, with the measurements of δ.

The failure to recognize the secondary maxima at the very beginning of one's experience with binaural beats, (7), can be explained by our assumption that the blending of the sensations produced by the B instruments is only partial.

The Localization. Our theory is in accord with the experimental fact that during a portion of the beat-cycle, *i. e.,* in the region $\epsilon = 0°$, the localization is external, and in the portion near the region $\epsilon = 180°$, the localization is internal. For in the former case the excitation (of the A instruments) is by the usual means, air-drum-skin-etc., and in the latter case, the excitation (of the B instruments) is through the skull which, as we know by direct experiments, (8), gives an internal localization.

In order to bring our theory into agreement with all the experiments on localization described above, we must make two additional assumptions, namely, that the localization with the A instruments depends chiefly upon phase difference and that with the B instruments chiefly upon intensities. The former assumption will perhaps not be regarded by all as a reasonable[28] one, but the latter involves a principle generally accepted.

At $\epsilon = 0°$, the resulting phase difference at the pair of A instruments is zero and hence the localization is in the median plane. The fork near to an ear will furnish the greater portion of the amplitude and thus the variation of phase at each A instrument can be assumed to be chiefly determined by the phase

[27] See article on "Theory of Binaural Beats," Physical Review, 1917.

[28] J. Peterson, loc. cit., assumes that the perception of phase differences is cortical in origin.

of the nearer fork. Thus the localization in the region, $\epsilon = 0°$ follows, approximately, the variation of ϵ. And we note that the movement of localization, experimentally, is in accord with this result, for it seems, as in (8), to proceed at a uniform angular velocity.

In the region $\epsilon = 180°$, the excitation of the B instruments is the more prominent and hence controls the sensation experienced. That the localization at or in an ear occurs both at $180° - \delta$ and $180° + \delta$, where δ is less than, say, 45° is in accord with theory in that λ (or δ) must be small because of the high velocity of conduction in the bones of the skull. As yet the quantitative agreement, if we accept possible values of the velocity of sound in the bones, is not very good. But the possible conducting paths are numerous and do not actually terminate on the opposite side of the head from the source, and hence a quantitative agreement should not be anticipated.

The fact that δ is less than 45° makes the angular velocity of localization apparently much more rapid, as in (8), in the region $\epsilon = 180°$.

The continuation of the localization in the region $\epsilon = 0°$ and the absence of the localization in one ear when the source on that side is made of much less intensity, as in (10), is in accord with the above theory, for in the first case the difference of phase is only slightly dependent upon the equality of intensity, and in the second the maximum of intensity would be blotted out by the partial blending.

The difference in distinctness of the localization, described in (4), and the most accurate judgment of phase difference at $= 0°$ are experimental facts which are also in accord with our theory.

These explanations bring our theory of binaural beats into apparent harmony with the experimental facts concerning localization recorded above in (c), (d), (2), (4), (5), (6), (8) and (10).

The fact that the rotation of localization is not found with all observers is doubtless to be explained upon the basis that

the localization is a "second-order effect" as compared with intensity observations. The theory as here presented does not explain why the rotation of localization in the region, $\epsilon = 0°$, is similar with all frequencies, the same difference in phase meaning the same displacement. This phenomenon may have a wholly physical basis, but that is not certain. Full consideration has not yet been given to this point.

Anatomical Considerations

The query arises, does the structure of the ear permit of the assumptions as made? It would seem that no serious objection arises if we assume the instrument A to be the organs of Corti in the ductus cochlearis, and the instrument B to be the analogous organs in the saccule or the utricle. These organs in the cochlea and the saccule and utricle, have a common origin. Herrick[29] states that some physiologists have thought that cochlear and vestibular nerve systems are not wholly distinct and that the sense organs in the saccule may also function as a sound perceptor. This uncertainty concerning the function of the organs in the utricle and saccule on the part of physiologists does not serve as an objection to our theory. Indeed, the fact that the complex phenomena herein described can be accounted for by the assumption of the existence of the B instruments ought to assist in deciding this uncertainty. The fact that there is a great difference in the central connections of the cochlear and vestibular nerves is in harmony with our contention that the instruments B give sensations which do not completely blend and which do not produce external localization.

But are the vestibular and cochlear organs mechanically different? Obviously they are, for the latter are placed upon a membrane, and are apparently excited by the motions of the fluid in the scala media, and the motions of the basilar membrane and the tectorial membrane. The vestibular organs are located at the points where the membranous labyrinth is fastened to the skull itself. Moreover, the movement of the vestibular fluid must be very small for the vestibular canals are closed.

[29] G. J. Herrick, Introduction to Neurology, pp. 201-202.

Thus there would seem to be a mechanical difference which corresponds admirably with our assumption that instrument A is excited via the vibration of the membrana tympani and that instrument B is excited by the skull. The only change in our assumption induced by the structure of the ear would be the possible insertion of the word "chiefly" after each word "excited" occurring in the preceding sentence.

A paper of importance has just come to the writer's notice. D. Richards[30] reports upon experiments with a guinea pig in which there were removed (1) both cochleae, (2) both cochleae, and one vestibule and (3) both labyrinths. His conclusion is that sound stimuli may be regarded as adequate stimuli for the vestibular apparatus, but that no conclusion can be drawn as to any sensation produced in this way. This contribution is entirely favorable to the theory presented in this paper.

Restatement of Theory and Agreement with Experiment

By the assumption of two physical instruments A and B on each side, having a difference in mechanical arrangement, a difference in the method of excitation, a difference in the blending of the sensations of the A pair and B pair of instruments and an independence of sensations produced by the two pairs and the additional assumptions that instruments A, the usual instruments of hearing, can perceive phase differences, and that in the skull conduction to the B instruments there is an additional phase change of 180°, we have a theory that explains fairly satisfactorily the presence of the primary maxima, the secondary maxima, the appreciable length of beat-period requisite for the appearance of the latter, the localization in the entire beat-cycle, the difference in the nature of the localization and in the quality of the sound in the two halves of the beat-cycle, the rapid movement of localization between the secondary maxima, the variation of the separation of the secondary maxima with the fork-frequency, and its independence of beat-frequency, the difference in the effects of unequal intensities upon the localization in the two halves of the beat-cycle, and finally the

[30] D. Richards, Zeitsch. f. Biol. 1916, 66, pp. 579-609.

simultaneous disappearance of the secondary maxima and the localization in the region of 180° phase-difference.

Our theory does not lead us to identify the B instruments, but merely to suggest the possibility of these instruments being located in the utricule and saccule. Anatomical and physiological considerations give evidence in favor of the suggestions.

Inasmuch as all of the experiments recited in the foregoing pages were performed with binaurals that were inserted into the ears, the query naturally arises as to whether or not the phenomena involved are dependent upon this insertion of binaurals. From a physical point of view the partial closing of the external meatus would not be likely to change the characteristics of the phenomena. But the actual experimental evidence of the truth of the similarity of results with inserted binaurals and binaurals free from the head is obtained by a comparison of the results of Rostosky,[31] who used the latter type of binaurals, and of the writer who tried both in all experiments excepting the quantitative measurement of δ. The results of these two investigators are in entire agreement. Hence it is believed that the theory is equally applicable to open-ear-binaurals.

The theory is, of course, incomplete, but it has herein proved to be very successful and hence promising. There are no conspicuous phenomena connected with binaural beats that are not in accord with the theory, but there are minor points that should receive further consideration. For example, S. P. Thompson[32] has found binaural beats occurring when the tones used have frequencies of almost two to one.

In the theory presented it is assumed that the sensation produced depends upon the intensity or upon the square of the amplitude of vibration. But this assumption is open to question. Might not the sensation vary with the absolute value of the amplitude? Might not an assumption be made that will permit of both the primary and secondary maxima without a necessity for the hypothesis of a second organ of hearing? The first question has been answered by a theoretical investigation

[31] Loc. cit.

[32] S. P. Thompson, Phil. mag. (5), IV, 1877, p. 274.

assuming the variation of the sensation to depend upon the absolute value of the amplitude. This investigation shows that the same maxima and minima occur as with the assumption of sensation as a function of intensity. But the second question remains unanswered. The writer merely opines that no assumption as to the variation of the sensation will lead to the possibility of the primary and secondary maxima without the additional assumption of two hearing instruments.

I wish to express my thanks to Mr. F. C. Bruene and to Dr. Harold Stiles for their assistance in different portions of the experimental work herein described.

PHYSICAL LABORATORY
STATE UNIVERSITY OF IOWA.

CORRELATION OF FACTORS IN MUSICAL TALENT AND TRAINING

BY

CARL E. SEASHORE AND GEORGE H. MOUNT

CONTENTS

Introduction
Sources of material
Tables of correlation
Comments on the correlations
Some general considerations and conclusions
References

INTRODUCTION

We have here undertaken to assemble and interpret, for the technical reader, some of the data available in this laboratory on the interrelations of various factors in musical talent.

For the last fifteen years or more, a mass of data bearing on this subject has been accumulating until it is now unquestionably the largest collection of its kind. Most of these data have been presented, or will be treated in the discussion of individual problems by different investigators; but the presence of this material suggests the possibility of a mass-survey of the various kinds of records in this series for the purpose of finding out something about the inter-relationships of the factors under consideration.

All the records considered have been taken upon University students in the elementary course in psychology, all in or above the sophomore year. The plan has been to take the entire class in order to eliminate the error of selection. The tests have been made under very favorable conditions, either as class tests under accurate control or as individual tests in the special measuring rooms of the psychological laboratory.

Some of the tests have been made for more than ten successive years on from one hundred to three or four hundred different students. Instead of attempting to use all of the data available, we have selected certain groups, or cross sections, which are representative, namely the records taken in the years '07, '09, '11,

and '16. This selection excludes a few relevant tests which were taken only in other years.

Dr. Mount originally undertook this task and completed his doctor's thesis on this subject in 1910, having developed the material given here for the years, 1907 and 1909. The fact that the same investigation was continued by his immediate followers led to the postponement of publication and, during the time that elapsed, certain parts of the tests particularly the section on the distribution of musical talents, have been superceded by more intensive studies along some of these lines, notably, the measurements on pitch by Smith (16), and on singing by Miles (6).

In Dr. Mount's original treatment of this subject, a new method of expressing correlations was developed consisting essentially of a characteristic figure, (K), representing the distribution of records into quartiles or octiles, a figure somewhat analagous to the "r" in the Pearson coefficient. Since this method has not been followed up in later researches, it was thought best to use the coefficient R of the Spearman foot-rule in the present article in order that these data might be more nearly comparable to those contained in the data worked over by his followers.

It would be poor economy for the reader to have Dr. Mount's original interpretation of the correlations presented separately from those of his followers. Therefore, with his courteous consent, the correlation data which he had worked out are here used as a part of the larger body of the same kind of material, and his original thesis will not be published for the reasons given; namely, that other parts of his thesis are superceded by later work now published, that the reading public will welcome the whole report as one unit, and that his original interpretations, in so far as they are substantiated, are embodied in the present interpretation of the material as a whole.

The group of data for the year, 1911, was collected by Dr. Rollin M. Stewart, who put much valuable time and energy into the evaluation of the questionnaire returns and the ranking of students on the questionnaire. He did not find time to work out the correlations and write up the subject himself, and therefore turned over the material to the writer.

The materials for 1916 were assembled from the various in-

vestigations by Mrs. Esther Allen Gaw to whom the writer is also indebted for continued help in the work in her capacity as research assistant in the field of this investigation. The first named of the authors has written this entire article on the material mentioned and takes all responsibility for the interpretations herein presented. He wishes to express his hearty appreciation of the coöperation of many whose names are not mentioned. He, and all the other workers in this field, are under special obligation to Dr. Mabel Clare Williams for her valuable guidance and supervision of these tests.

The terse account of the various tests which follows can be excused only on the ground that in most cases reference is given to published accounts which should be read in detail if the procedure is to be fully understood or repeated.

It would not seem necessary to go into details as to the conditions of the experiments, which would in themselves fill a large volume, because our only interest in the present issue is in the correlation problem; it is sufficient to know that the same procedure was followed for all the members of a given group or year for which the figures have been compared.

Likewise one might go into great detail in the discussion of the merits and demerits of various methods of computing and representing correlations, means of attenuating, weighting, etc., but that we must also forego and proceed by the simplest method. It should be said that the demand for considering attenuations of correlations and other adjustments was reduced to a minimum by the fact that the measurements were made under good laboratory control, and that the number of cases on which each correlation was based is relatively large. In '07 the minimum was about 190, in '09 about 200, in '11 about 225, and in '16 about 225, but each year a much larger number of cases was taken for some of the correlations.

As stated above the Spearman foot-rule was used for '07 and '09 because much of the material could be best handled by the method of ranking. This was also used for the questionnaire returns in '11. Otherwise, in '11 and '16 the coefficient "r" of the Pearson product-moments method was used. This diversity of standard is unfortunate and must be kept in mind in reading.

The Spearman method, larger formula, gives the higher correlation as is shown by comparing the "R" and the "r" in two actual items.

'07 R, .40, p.e. .04 is equivalent to r, .59, p.e. .03

'09 R, .49, p.e. .03 is equivalent to r, .70, p.e. .02

The general formula by which one may be converted into the other in a very rough estimate is either $r = \sin\ (\frac{\pi}{2} R)$ which is Spearman's empirical formula, or $r = 2 \cos\ \frac{\pi}{3} (1\text{-}R)\text{-}1$ which is Pearson's theoretical formula to give the r of the product-moments method from Spearman's R.

SOURCES OF MATERIAL

1. *Pitch discrimination.*—This measurement was made according to the standard conditions laid down by the writer in the article on the standardization of measurement of pitch discrimination (12). Briefly, the standard was 435 v.d. and ten differential forks ranged from ½ v.d. approximately in a geometric ratio of the second order up to 30 v.d. The tones were sounded through Helmholtz resonators. The "heterogeneous" method, described at length in the above reference (12), was followed, and it was found possible to make from 170 to 200 trials in the one hour.

These tests were made on groups of students from 100 to 200. In '09 this test was followed by individual tests, using the method of constant stimuli at the threshold found in the group test. The records for '09 were accordingly labeled "group tests" and "individual tests" respectively, and the plain "P.D." in Table II, without qualification, denotes the mean between the two.

In '09 a practice series of three periods was also given. The observers were divided into three divisions on the basis of records made in the first test. Division A included all whose threshold was below 4, Division B those whose threshold was between 4 and 10, and Division C those whose threshold was above 10. The three divisions were tested simultaneously but in different rooms by different experimenters. Division A was given fifty trials upon each of the increments 5, 3, 2, and 1 respectively. Division C was tested in the same way upon increments 12, 8, 5, and 3, respectively, and Division C was tested upon increments

30, 17, 13, and 12 respectively. The threshold was determined by applying the Fullerton-Cattell formula (3) with reference to the cases in which the records were between 65 and 85 per cent.; records on other intervals were discarded except when new intervals fell within the limits indicated.

Measurements in 1907 were made by the writer and Dr. Williams, in 1909 by Dr. Mount, in 1911 by Dr. Smith (16) and in 1916 by Mrs. Gaw.

2. *Consonance.*—All the measurements on consonance were made by the method described in Dr. Malmberg's article on the perception of consonance in this volume (5). Briefly, the procedure was this: Malmberg had ranged all possible two-clangs in the octave from c′ to c″, in the order of consonance-dissonance, for the piano. This order was used as a norm and the measurement was made by the method of paired comparison, each two-clang being compared with every other as sounded on a well tuned piano. Errors were then checked by the norm named above, and these errors were weighted according to the magnitude of displacement from the true order according to the norm. Thus if c′d′ was said to be better than c′c″ a demerit of 11 was given. If c′d′ was said to be better than c′g′ a demerit of 9 was given, etc., there being a scale of demerits from 11 to 1 according to the magnitude of the error. The record was kept in terms of the per cent. of weighted credits for right judgments above the one-half which would be right by chance. For full understanding of this method it is necessary to read Malmberg's original statement.

A critical feature of this test is the definition of consonance. It was pointed out to the class that consonance may be regarded from two points of view, that of feeling and that of knowing, and that our point of view was the latter, and consonance was accordingly defined by the following placard:

1. Blending: a seeming to agree, to belong together, to be relatively pure.
2. Smoothness: relative freedom from beats.
3. Purity: thinness of tone.

This was read and interpreted to mean, "When the two tones of a two-clang tend to blend or fuse and produce a relatively smooth and pure resultant, they are said to be consonant."

The '11 records were taken by Dr. Malmberg and the '16 records by Mrs. Gaw. Attention should be called to the fact that this test has since then been modified as reported by Mrs. Gaw in this volume (4).

3. *Intensity discrimination.*—The measurement of discrimination for intensity of sound was made in the light-, sound-, and jar-proof room (7) of the psychological laboratory with the audiometer (10). The interrupted current reduced to 1.8 volts from a 100 v.d. tuning fork was sent through the audiometer. This produced a tone which was clearly audible. The strength of this tone was kept constant.

The '11 tests were made by Mr. W. S. Smiley and the '16 by Miss Marie M. Agnew.

4. *Hearing ability.*—Hearing ability was also measured by means of the audiometer (10) in the light-, sound-, and jar-proof room (7).

The audiometer was used, as above, but without the tuning fork, and the current was reduced to the standard strength according to the audiometer galvanometer. Ten determinations were made of the upper limit and ten of the lower limit of the threshold and the mean between the averages of these two was taken as the characteristic measure of hearing ability.

All known subjective and objective sources of error were under reasonable control. The record was taken for the better ear according to the judgment of the observer.

These tests were made by Mr. W. S. Smiley in '09 and by Miss Marie M. Agnew in '16.

5. *Time-sense.*—The measure of time-sense was the least perceptible difference in the length of two empty intervals marked off by three clicks in a telephone receiver. In 1911 the timing was done by the synchronous motor as described by Ross (9). An adjustable contact lever was attached to the shaft of the motor in such a way that when the motor revolved at the rate of one revolution per second, this contact could be shifted from the standard point, for one of the intervals, making it shorter than one second by a given differential step. The steps employed were 1, 2, 3, 5, 8, 12, and 17 hundredths of a second.

In 1916 the same measurement was made by means of the

Dunlap pendulum (2) which is a very large and accurately built seconds pendulum. In this case the middle click, *i.e.*, the click at the end of the first interval, was placed earlier or later than the standard according to the same differential steps as before. This resulted in the lengthening of one interval and the shortening of the other whereas in the '11 test with the synchronous motor one interval was kept constant and the other shortened. This difference was, however, of no consequence for the present purpose. The pendulum was substituted for the synchronous motor in order to develop a norm for the pendulum which can be used in simple and readily transportable form.

The '11 tests were made by Mr. Ross (9) and the '16 by Miss Pauline Peters.

6. *Motor ability.*—The test of motor ability was the customary procedure of tapping with a telegraph key adjusted with a spring requiring about 50 grams of pressure and moving through a distance of one millimeter. The record was made with a multiple recorder tracing on ticker tape by the method described for the psychergograph (9) by the writer. This multiple recorder is a sort of a "universal apparatus" for the laboratory, convenient for all sorts of prolonged time and signal records. The time line was traced from a 25 v.d. tuning fork. When indelible leads are used and the tape is run under a moist sponge, the moisture of the paper develops the light tracing into a clear and permanent record.

This and the next three tests were made by Miss Pauline Peters in '16.

7. *Free rhythm.*—Free rhythm was measured in the terms of the average error in the free marking of time at the suggested rate of approximately one beat per second. The record was made on the multiple recorder as above.

We have followed custom in calling this rhythm. It is really not a rhythm test but a test of motor-time, or time-action. The same is true of "regulated rhythm" also.

This and the next test on regulated rhythm were made by Miss Margaret Cummings in '11.

8. *Regulated rhythm.*—Regulated rhythm was measured in terms of the average error in keeping time with a click produced

at intervals of one second in a telephone receiver operated through a seconds pendulum. The original object of this experiment was to study personal equations but for the present purpose only the average error was counted.

9. *Rhythmic judgment.*—"Rhythmic judgment" is something of a misnomer. It consisted in marking time with the telegraph key through the multiple recorder (13) as before, the instructions being to mark time, giving in each measure two short and one long intervals. The record was made in terms of the deviation of the long interval from the sum of the two short.

10. *Singing the keynote.*—All measurements on the singing of pitch were made by means of the tonoscope (14). Singing the keynote was measured in terms of the average error in reproducing the standard tone sounded by a tuning fork before a resonator. In most cases the standard was 256 v.d., the women being allowed to sing this pitch and the men an octave lower. This method was followed because the men preferred to take the key from a 256 v.d. fork rather than from a 128 v.d. fork, as a result of association with notes on musical instruments. This measurement was not made in and by itself but was taken from the singing of the keynote in singing intervals, scales or minimal change; but for each year it was the same for the entire group, which is all that is necessary for us to ascertain for the present purpose. It will be observed that the note sung was in the middle register of voice both for men and women.

All the measurements of singing were made by Dr. David Allen Anderson (1) in '09, by Dr. Walter R. Miles (6) in '11, and by Esther Allen Gaw in '16.

11. *Singing the interval.*—In '09 the test consisted of the singing of the three intervals, the major third, the fifth, and the octave after sounding the keynote in each trial, the measure being in terms of the average error, as calculated from the standard as sung rather than as sounded by the keynote.

In '16 this test consisted of the singing of the first two phrases of America, record being made of the average error in the sounding of the notes which occur for the words, *my, 'tis, sweet, of, of, sing.*

12. *Singing the scale.*—Each observer was required to sing

the natural scale within a middle register of the voice, in just intonation.

13. *Voluntary control of pitch.*—This is the measurement of the least producible difference, or the minimal change, in sharp or flat as described by Miles (6).

14. *Singing.*—This ranking is based on the mean between the singing of the key and the minimal change as described above, a single measure being used for the two.

15. *Auditory imagery.*—In '16 the test as outlined in the writer's Manual of Elementary Experiments in Psychology (11, Ch. IX), was supplemented by a one-hour experiment devoted entirely to auditory and motor imagery in which the stimuli were a broken melody, printed words, spoken words, and sounds to be imaginally imitated. The records used for the purpose of correlation were the mean between the Manual test and the record for this special test on auditory imagery. The method of ranking the images described in the Manual was followed and the records were kept in terms of the percentage of points possible.

The '11 tests were made by Dr. D. J. Macdonald, and the '16 tests of this and motor imagery by Miss Marie Agnew.

16. *Motor imagery.*—Motor imagery was measured according to the directions given in the Manual of Elementary Experiments in Psychology (11, Ch. IX), section on motor imagery. This was supplemented by a test on the imaging of a series of tones. The mean between these two records was used as the final grade.

In this test, although the observer was warned to distinguish between motor sensations and motor imagery, it is safe to assume that this was not done successfully; and it is probable that much of what the untrained observer did record as motor imagery may have involved motor sensations.

17. *Tonal memory.*—Tonal memory was measured in terms of the memory span for notes on the piano according to the method developed by Sodergren (17). Briefly, a series of notes in the middle octave of the piano was sounded. Immediately thereafter the same series, with but one change, was sounded and the observer was required to tell which of the notes was different in the second stimulus group. The spans of two notes, three notes,

four notes, five notes, and six notes were used under controlled conditions. The test lasted one hour and the record was made in terms of the percentage of right cases in all the trials. From these records an evaluation was worked out showing to what a given percentage of right cases in the total number of trials would be equivalent in terms of 75 per cent. of right cases on a given span.

This test was made in '16 by Mrs. Gaw.

18. *Supplementary information.*—All persons examined were requested to fill out the following questionnaire:

Name..... Age.....
Date..... Hour..... Class.....

Please give as specific and detailed information as possible in regard to your:

I. Musical Training
   1. In public schools
   2. Private vocal lessons (when, where, how long, etc.)
   3. Private instrumental lessons (when, where, how long, etc.)

II. Musical Environment
   1. Instruments in your home, and their use
   2. Musical encouragement at home (trained voices in family)
   3. Opportunities for hearing music of any sort (specific)

III. Musical Expression
   1. Favorite selections you can sing (by ear? by note?)
   2. Favorite selections you can play (by ear? by note? instruments?)
   3. Singing or playing in public (parts, occasions, etc.)

IV. Enjoyment of Music (What Actually Appeals to You)
   1. Vocal (solo, quartette, chorus, opera, popular songs, classics, religious, secular)
   2. Instrumental (solo, symphony, band, etc.)
   3. Characteristic effects of music (mental, physical)

All were charged to make the information as full, specific, and concrete as possible. Often a questionnaire was returned for more accurate information. In many cases this information was supplemented by personal interviews on significant points. Notes were also taken in the various tests placing on record facts which could be used in this section.

All the ratings in the section were made without knowledge of the objective records. The rating is arbitrary, but was, nevertheless, done with extreme care and patience in the effort to sift the facts. Various devices were used for the evaluation of answers in the different years. The record for musical training being made in terms of hours of instruction, public school train-

ing had to be reduced to this unit. Vocal and instrumental lessons were treated with different value.

The most elaborate evaluation is that worked out by Dr. Rollin M. Stewart for '11, but it is not necessary to go into details about this procedure because the same method was followed for all in the entire group for a given year and the title of each item is perhaps sufficiently descriptive.

Mount differentiated instrumental training from vocal training for the purpose of tracing the relationship of each of these two factors separately. The first item on the questionnaire, musical training, is, therefore, based on a composite record of instrumental and vocal training. The same principle of division was followed for "musical expression."

TABLES OF CORRELATION

The records for the four years selected are presented in Tables I to IV. Attention is called to the fact that the first two and a part of the third are in terms of the Spearman R; and that the third, in part, and the fourth wholly are in terms of the Pearson r. Absence of sign indicates a positive correlation.

TABLE I—*Series of 1907*

	P.D.	M.Tr.	I.Tr.	V.Tr.	M.Env.	M.Ex.	I.Ex.	V.Ex.
Musical Training	.31 .04							
Instrum. Training	.27 .04	.94 .00						
Vocal Training	.20 .05	.34 .04	.23 .05					
Musical Env.	.12 .05	.26 .05	.25 .05	.15 .05				
Musical Exp.	.38 .04	.51 .03	.42 .03	.25 .05	.35 .04			
Instrum. Exp.	.28 .04	.65 .02	.68 .02	.12 .05	.38 .04	.77 .02		
Vocal Exp.	.32 .04	.14 .05	.06 .05	.24 .05	.21 .05	.61 .03	.23 .05	
Musical Enj.	.27 .04	.22 .05	.23 .05	—.04 .05	.03 .05	.12 .05	.13 .05	.08 .05

The upper figure in each pair is R of the Spearman foot-rule formula; the lower is the p. e. of R.

TABLE II—*Series of 1909*

	P.D.	P.D. Av.	P.D. Ind.	S.K.	S.I.	S.Sc.	S.	M.T.	I.T.	V.T.	M.Env.	M.Ex.	I.E.	V.E.
P.D.	.64													
Av.	.02													
P. D.	.68	.68												
Ind.	.02	.02												
Singing	.14	.32	.21											
Keynote	.05	.04	.04											
Singing	.15	.28	.19	.45										
Interval	.05	.04	.05	.04										
Singing	.24	.34	.25	.42	.47									
Scale	.04	.04	.04	.04	.03									
Singing	.26	.37	.29	.76	.75	.78								
	.04	.04	.04	.02	.02	.01								
Musical	.16	.19	.10	.14	.23	.18	.18							
Training	.04	.04	.05	.05	.04	.05	.05							
Instr.	.15	.16	.11	.11	.39	.14	.18	.96						
Training	.05	.05	.05	.05	.04	.05	.05	.00						
Vocal	.18	.31	.20	.26	.15	.07	.17	.28	.13					
Training	.04	.04	.04	.04	.05	.05	.05	.04	.05					
Musical	.17	.28	.17	.24	.18	.16	.22	.30	.28	.12				
Env.	.04	.04	.04	.04	.05	.05	.04	.04	.04	.05				
Musical	.31	.42	.36	.23	.24	.31	.32	.48	.47	.26	.30			
Exp.	.04	.04	.04	.04	.04	.04	.04	.03	.03	.04	.04			
Instr.	.22	.29	.29	.15	.21	.25	.32	.60	.66	.14	.34	.79		
Exp.	.04	.04	.04	.05	.04	.04	.04	.03	.02	.05	.04	.01		
Vocal	.22	.31	.27	.24	.28	.34	.34	.07	.00	.54	.11	.56	.18	
Exp.	.04	.05	.04	.04	.04	.04	.04	.05	.05	.03	.05	.03	.04	
Musical	.13	.12	.07	.04	.05	.06	.01	.03	.05	.19	.15	.16	.20	.19
Enj.	.05	.05	.05	.05	.05	.05	.05	.05	.05	.04	.05	.04	.04	.04

The upper figure in each pair is R of the Spearman foot-rule formula; the lower is the p. e. of R.

TABLE III—*Series of 1911*

	P.D.	C.D.	I.D.	F.R.	R.R.	S.	A.I.	M.I.	M.T.	M.Env.
Cons.	.33									
Discr.	.05									
Intensity	.13	—.03								
Discr.	.05	.05								
Free	.27	.03	—.05							
Rhythm	.05	.05	.05							
Regulated	.01	—.11	.18	.04						
Rhythm	.05	.05	.05	.05						
Singing	.12	.04	.14	.26	.07					
(Ave.)	.05	.05	.05	.05	.05					
Auditory	.29	.14	.05	.22	.17	.07				
Imagery	.04	.05	.05	.05	.05	.05				
Motor	.02	—.02	—.05	.06	.11	—.07	.52			
Imagery	.05	.05	.05	.05	.05	.05	.04			
Musical	.21	.14	.03	.06	.03	.09	.21	.04		
Training	.03	.03	.03	.03	.03	.03	.03	.03		
Musical	.13	.03	.01	.09	.07	.05	.13	—.03	.28	
Env.	.03	.03	.03	.03	.03	.03	.03	.03	.03	
Musical	.06	—.02	.01	.04	—.01	.06	.11	.11	.10	.13
Enj.	.03	.03	.03	.03	.03	.03	.03	.03	.03	.03

The upper figure is the r of the Pearson product-moments method; except in the last three lines, for musical training, musical environment, and musical enjoyment, which represent the Spearman R; the lower is its p. e.

The upper figure in each pair is R of the Spearman foot-rule fo

TABLE IV—*Series of 1916*

	P.D.	C.D.	I.D.	H.A.	T.S.	M.A.	F.R.	R.R.	R.J.	S.K.	S.I.	V.C.	A.I.	M.I.
Cons.	.21													
Discr.	.04													
Intensity	.09	.11												
Discr.	.04	.04												
Hearing	.12	.01	.16											
Ability	.04	.04	.04											
Time	.17	.07	.15	.03										
Sense	.04	.05	.04	.04										
Motor	—.02	.03	.01	.24	.01									
Ability	.05	.04	.04	.04	.04									
Free	.18	.03	.17	.005	.08	.09								
Rhythm	.04	.04	.04	.04	.04	.04								
Regulated	.21	.01	.05	.09	—.01	.12	.14							
Rhythm	.04	.04	.04	.04	.04	.04	.04							
Rhythmic	.22	.30	.001	—.06	—.03	.03	.15	.00						
Judgment	.04	.04	.04	.04	.04	.04	.03	.00						
Singing	.21	.30	.28	.11	.001	.10	.01	.13	—.04					
Keynote	.04	.04	.04	.04	.04	.04	.04	.04	.04					
Singing	.22	.21	.27	.32	.10	.21	.07	.23	—.05	.61				
Interval	.04	.04	.04	.04	.04	.04	.04	.04	.04	.03				
Voluntary	.38	.01	.10	.001	.01	—.05	.07	.11	.08	.55	.32			
Control	.04	.04	.04	.04	.04	.04	.04	.04	.04	.03	.04			
Auditory	.29	.06	.04	.03	.19	.06	—.01	.001	.08	.24	.40	.29		
Imagery	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04		
Motor	.07	.01	.01	.22	.13	.10	.07	.08	.07	.19	.21	.08	.42	
Imagery	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	
Tonal	.52	.23	.26	—.01	.17	.10	.09	.05	.22	.32	.30	.50	.40	.16
Memory	.03	.04	.04	.04	.03	.04	.04	.04	.04	.04	.04	.03	.04	.04

The upper figure is the r of the Pearson product-moments method; the lower is its p. e.

COMMENT ON THE CORRELATIONS

Consonance and pitch discrimination: '11, r, .33, p.e., .05; '16, r, .21, p.e., .04.

Although the perception of consonance depends upon pitch discrimination it is not closely limited by that, for it is safe to say that eighty-five per cent. of all persons have sufficiently good pitch discrimination to hear consonance well enough to secure a record above the average in so far as it depends upon pitch discrimination alone. In other words, pitch discrimination is not the only determining factor in consonance; blending, fusion, smoothness, and purity may be perceived to some extent apart from perception of pitch, as such.

Intensity discrimination and pitch discrimination: '11, r, .13, p.e., .05; '16, r, .09, p.e., .04.

These two factors have but little in common. What there is may perhaps be thought of in general as a tendency toward ear-mindedness. Physiologically the two factors are quite independent in normal hearing.

Hearing ability and pitch discrimination: '16, r, .12, p.e., .04.

It is conceivable that defective hearing-ability (acuity) may be due to anatomical and physiological conditions which involve the organ of Corti and would therefore result in defective pitch discrimination; but most of the causes of defective hearing are due to disturbances elsewhere in the ear. The correlation may be due in part to the pathological condition in the ear, and in part to factors in psychological disposition, such as ear-mindedness and power of application.

Time-sense and pitch discrimination: '16, r, .17, p.e., .04.

Pitch and time are rightly regarded as independent variables in hearing. The variability in pitch is due mainly to structural or functional differences in the ear, whereas variability in the perception of time is due to motor and central factors. It may be that a part of the correlation is due to the capacity for improving the cognitive threshold with practice, in cases of relatively low ear-mindedness; *i.e.*, some, who are dull and are not ear-minded, may have gotten low records in both tests because of their general inability to respond to the test.

Motor ability and pitch discrimination: '16, r, -.02, p.e., .05.

This tallies well with what we know of the independence of these factors; and the absence of correlation speaks well for the elemental character of these two tests. That is, each is so simple that it does not involve acquired skill or general intelligence on a large scale. The two tests were equally fair to the bright and the dull, the quick and the slow.

Free rhythm and pitch discrimination: '11, r, .27, p.e., .05; '16, r, .18, p.e., .04.

Free rhythm involves all the elements of time-sense plus musical feeling and a certain agility in action; it is the motor aspect of which time-sense is the sensory or central. The tests in free rhythm for '11 were really more reliable than those for '16. The correlation, it is reasonable to suppose, may be traced to the feeling for, or interest in, sounds which is favored by a keen ear for pitch; that is, hearing differences of pitch tends to result in expression of pitch, and all musical expression tends to become rhythmical.

Regulated rhythm and pitch discrimination: '11, r, .01, p.e., .05; '16, r, .21, p.e., .04.

Probably a mean between these two records would represent the facts in the long run. Regulated rhythm is time-sense plus precision of action. The feeling element of free rhythm is not so conspicuous. Regulated rhythm represents a higher cognitive or central affair, with temporal precision in action, and is quite independent of sensory limitations.

Rhythmic judgment and pitch discrimination: '16, r, .22, p.e., .04.

Rhythmic judgment does not involve the feeling element and here again the correlation may be laid to that factor, as in free rhythm.

Singing the keynote and pitch discrimination: '09, R, .14, p.e., .05; also R, .32, p.e., .04; '16, r, .24, p.e., .04.

Miles (6, p. 61) found a correlation between these two factors as follows; for eighty-four men, the correlation between pitch discrimination and the average error in singing the keynote was r, .04, p.e., .07; and for pitch discrimination and the

constant error in singing the keynote it was r, .08, p.e., .07. For one hundred and four women he found, for pitch discrimination and the average error, r, .27, p.e., .06; and, for pitch discrimination and the constant error, r, .11, p.e., .06. These facts of the comparatively low correlation between the hearing of pitch and singing of pitch may be somewhat surprising, but the singing test is far from elemental; and, as in consonance. the average pitch hearing is good enough for excellent pitch singing. Singing, which is motor control, often does not reach the hearing limit and in so far as it does not, there should be no correlation.

On this question of relation between hearing of pitch and singing of pitch in learning to sing, an investigation has been completed by Mr. Carl J. Knock and will be published in the next volume of these *Studies.*

Singing intervals and pitch discrimination: '09, R, .15, p.e., .05; also R, .28, p.e., .04; and R, .19, p.e., .05; '16, r, .32, p.e., .04.

Miles (6, p. 61) found, for forty-eight men, a correlation of R, .33, p.e., .08 for pitch discrimination and average error in singing the interval, and for pitch dicrimination and the constant error in singing the interval, r, .15, p.e., .07. For one hundred and four women the corresponding records were r, .51, p.e., .05, and R, -.07, p.e., .06 respectively. These strong correlations between pitch discrimination and the average error in singing the interval cannot be attributed wholly to the presence of pitch discrimination in the hearing of intervals. The common element, that is, a general sense of pitch must be taken into account. Intervals must be learned; those who have a good ear for pitch are the ones who practice and develop an interest in intervals. The correlation was, therefore, due largely to the fact that those who have a good sense of pitch take time to practice on the intervals.

The records by Miles on correlation of pitch with constant error in singing involve different factors. The constant error is due to certain motives for illusion which become fixed and habitual; by some motive for constant error, a person may have a concept of the interval which is persistently flat or sharp

by a certain amount. The constant error is in part an error of concept and in part an error of memory for that concept. This accounts for the low correlation between pitch discrimination and constant error in the interval found by Miles.

Singing the scale and pitch discrimination: '09, R, .24, p.e., .04; also R, .34, p.e., .04; and R .25 p.e., .04.

What has been said about the average error in singing of intervals applies equally here.

Voluntary control and pitch discrimination: '16, r, .38, p.e, .04.

Of all types of singing this is the most closely dependent upon pitch discrimination because the problem is to sing the least producible difference. All things taken into account, the correlation here given seems to be characteristic of this relationship. It must be remembered that, without practice, one is not likely to learn the motor control that will actually produce as small a difference as can be heard.

Singing and pitch discrimination: '09, R, .26, p.e., .04; also R, .37, p.e., .04; also R .29, p.e., .04; '11, r, .12, p.e., .05.

These composite records are not of much value for correlation, because different principles are involved in each of the component elements.

Auditory imagery and pitch discrimination: '11, r, .29, p.e., .04; '16, r, .29, p.e., .04.

Where there is a common element which relates pitch discrimination to other sensory and motor processes in singing, it probably comes most frequently through the channel of auditory imagery. The correlations here shown are good. One may question which of the two factors is the basic. Does auditory imagery depend upon pitch discrimination in so far as pitch is concerned, or does the opposite relationship obtain? Auditory imagery, which we ordinarily designate as ear-mindedness is, as it were, a general vehicle for other factors. Without entering into the dispute as to the rôle of images in successive comparisons of tones, we may assert that the image is present and probably favors discrimination. The correlation here shown seems to represent fairly the judgment of most observers

in regard to this relationship according to introspections. Those who have low auditory imagery tend to have poor and uncertain pitch discrimination.

Motor imagery and pitch discrimination: '11, r, .02, p.e., .05; '16, r, .07, p.e., .04.

From the fact that a keen sense of pitch would naturally develop motor responses and images of such responses to sound, one might have expected a larger correlation here. Its absence is conspicuous; it is undoubtedly due in part to the difficulty untrained observers find in identifying motor images.

Tonal memory and pitch discrimination: '16, r, .52, p.e., .03.

Memory is notably a faculty that responds to training. The accurate hearing of tones leads to the development of a natural memory for tones; conversely, indifferent hearing is even more certain to suppress attention to tones. Indirectly also, those who had formed the habit of hearing pitch were favored in the detection of the changed note which was often only a half-tone.

Musical training and pitch discrimination: '07, R, .31, p.e., .04; '09, R, .16, p.e., .04; '11, R, .21, p.e., .03.

a. Instrumental: '07, R, .27, p.e., .04; '09, R, .15, p.e., .05.

b. Vocal: '07, R, .20, p.e., .05; '09, R, .18, p.e., .04.

This item is perhaps the most significant. It raises two questions: Do those who have good sense of pitch seek musical training? And, does musical training improve pitch discrimination?

Observation of the many thousand cases we have, bearing on this point, leads us to the conclusion that those who have a good sense of pitch tend to like music, but only a comparatively small per cent. of these receive training. This fact in itself is perhaps adequate to account for the correlation here shown. The second question has been answered in the negative by Smith (16) and others, in so far as it refers to actual psycho-physic capacity.

It was thought that the relationship might be significantly different for instrumental and vocal music, the theory being that singing, like violin playing, offered special training in the precision of pitch hearing; but the figures do not warrant any such distinction. The explanation of this probably lies in the

fact that very little, if any, voice training is directed toward precision in pitch because the ear of both pupil and teacher is lenient and there is no objective aid employed as a check.

Musical environment and pitch discrimination: '07, R, .12. p.e., .05; '09, R, .17, p.e., .04; '11, R, .13, p.e., .03.

Environment is a part of education. It is also a mark of hereditary relationship. The relatively small correlation here shown for these two important factors supports the view that training in the hearing of pitch is a universal affair of normal hearing—like the perception of distance in vision. It reaches its normal development early in childhood through the ordinary use of sounds. If we may assume that the sense of pitch is an heritable character, that may account practically for the correlation here found.

Musical expression and pitch discrimination: '07, R, .38, p.e., .04; '09, R, .31, p.e., .04.

a. Instrumental: '07, R, .28, p.e., .04; '09, R, .22, p.e., .04.

b. Vocal: '07, R, .32, p.e., .04; '09, R, .22, p.e., .04.

These correlations involve essentially the same principles as in musical training. This concerns only the amount of musical activity. It is evident that the next point of interest should be a study of correlations with quality of achievement in musical expression.

The separation of vocal and instrumental expression seems to throw no light on the situation.

Musical enjoyment and pitch discrimination: '07, R, .27, p.e., .04; '09, R, .13, p.e., .05; '11, R, .06, p.e., .03.

One noteworthy peculiarity of special talent is that its absence is often not noted. One who has poor sense of pitch does not hear his own faulty pitch or the errors of others; he is satisfied and enjoys the many other aspects of music, apparently satisfied that music is to him what it is to others.

Average pitch discrimination in the practice series and the usual pitch discrimination, i.e., the first group test: R, '09, .64, p.e., .02.

Individual test of pitch discrimination by the method of con-

stant stimuli and the usual pitch discrimination of the first group test: R, '09, .68, p.e., .02.

Average of the pitch discrimination in the practice series and the individual test: R, '09, .68, p.e., .02.

In judging these relations it should be born in mind that a single group test cannot be counted on to establish the approximate physiological threshold in much more than half of the cases. Smith says: "Success in making a true measurement on an inexperienced observer in a single sitting varies with the knowledge, keenness, and care of the observer and the many objectively favorable or unfavorable conditions of the test as well as the experimenter; but, everything taken into account, it is safe to say that when an individual test is made under favorable conditions the approximate physiological threshold may be reached in a single sitting of less than an hour in more than half of the cases of adults or children who are bright and are old enough to understand the test. Even in group tests by the heterogeneous method one may reach in an hour the approximate physiological threshold of nearly half of the observers who are old enough and bright enough to observe." (16, p. 100).

It is, therefore, to be expected that the practice series and the individual test should both lead to readjustment. For this reason the correlation here shown is satisfactory as evidence of the stability of a single rough test. The fact that the correlation between the first group test and the individual test (presumably more accurate) is no greater than between the individual test and the average for the practice series, also speaks well for the first group test.

Intensity discrimination and consonance: '11, r, .03, p.e., .05; '16, r, .11, p.e., .04.

These two tests involve, to some extent, the same general power of critical attention to sound. It is difficult to explain the small correlation in '11.

Hearing ability and consonance: '16, r, .01, p.e., .04.

Time-sense and consonance: '16, r, '07, p.e., .05.

In these two cases it is rather remarkable that no significant correlation comes out here on account of the common element

of musical appreciation. The absence of correlation speaks for the elemental character of the tests.

Motor ability and consonance: '16, r, .03, p.e., .04.

No relation.

Free rhythm and consonance: '11, .03, p.e., .05; '16, r, -.03, p.e., .04.

Regulated rhythm and consonance: '11, r, -.11, p.e., .05; '16, r, .01, p.e., .04.

Comment for these two here same as for time-sense.

Rhythmic judgment and consonance: '16, r, .30, p.e., .04.

Singing the keynote and consonance: '16, r, .30, p.e., .04.

Singing intervals and consonance: '16, r, .21, p.e., .04

In the last three items the element of selection undoubtedly enters because those who have good consonance are likely to develop rhythmic judgment and accuracy in singing. It is difficult to identify other factors.

Voluntary control of pitch and consonance: '16, r, .01, p.e., .04.

Singing and consonance: '11, r, .04, p.e., .05.

The element of selection does not enter because these tests involve activities which never enter into music in this form, although it is the best test of pitch control. Voluntary control rests primarily upon pitch discrimination and motor ability. The absence of correlation with general singing ability in '11 is due to the fact that this record is based partly upon voluntary control.

Auditory imagery and consonance: '11, r, .14, p.e., 05; '16, r, .06, p.e., .04.

It seems strange that there is no higher correlation between these two factors, because the appreciation of consonance seems to rest largely upon that sort of mastery of details in tones which is ordinarily thought of as associated with the vivid representation of sounds.

Motor imagery and consonace: '11, r, -.02, p.e., .05; '16, r, .01, p.e., .04.

No correlation.

Tonal memory and consonance: '16, r, .23, p.e., .04.

This undoubted correlation is probably due in part to the

common element of familiarity with tones and interest in them.

Musical training and consonance: '11, R, .14, p.e., .03.

It is a comment on our processes of selection for musical training that there is so small a correlation between the sense of consonance and the opportunities for musical training.

Musical enjoyment and consonance: '11, R, -.02, p.e., .03.

No correlation; which indicates that the sense of consonance, like pitch discrimination, is an inborn capacity manifesting itself in, and developing through, the interest in, and dependence upon, the perception of qualitative differences in the world of sound which constitutes a part of the natural and ever present environment from early childhood, quite apart from its specialized form in music.

Musical enjoyment and consonance: '11, R, -.02, p.e., .03.

No correlation. This is a great puzzle. The evaluation of musical enjoyment is of course rather loose, while the measure of consonance is quantitatively quite accurate. It would be interesting to know to what extent this absence of correlation is due to inability of non-musical people to recognize the fact that they do not enjoy music as musical people do.

Hearing ability and intensity discrimination: '16, r, .16, p.e., 04.

Although these two activities seem similar they are not closely related, because hearing ability varies with the physical condition of the ear whereas, so far as we know, variation in the intensity discrimination for sound is largely due to psychological factors. It is however conceivable that some physical defects of the ear may be of such a nature as to affect both to the extent indicated by this correlation.

Time-sense and intensity discrimination: '16, r, .15, p.e., .04.

The common element here would seem to be the general capacity for sensory observation, if such there be.

Motor ability and intensity discrimination: '16, r, .01, p.e., .04.

Free rhythm and intensity discrimination: '11, r, -.05, p.e., .05; '16, r, .17, p.e., .04.

Regulated rhythm and intensity discrimination: '11, r, .18 p.e., .05; '16, r, .05, p.e., .04.

These three tests are alike in stressing precision.

Rhythmic judgment and intensity discrimination: '16, r, .00, p.e., .04.

Absence of correlation may be due to imperfection in the test of the judgment.

Singing the keynote and intensity discrimination: '16, r, .28, p.e., .04.

Singing intervals and intensity discrimination: '16, r, .27, p.e., .04.

Voluntary control and intensity discrimination: '16, r, .10, p.e., .04.

Singing and intensity discrimination: '11, r, .14, p.e., .05.

These are large correlations in the absence of dependence of one factor upon the other. The demand for auditory precision is the common element.

Auditory imagery and intensity discrimination: '11, r, .05, p.e., .05; '16, r, .04, p.e., .04.

Motor imagery and intensity discrimination: '11, r, -.05, p.e., .05; '16, r, .01, p.e., .04.

No correlation.

Tonal memory and intensity discrimination: '16, r, .26, p.e., .04.

From one point of view, intensity discrimination and memory span belong to the same type of process, differing only in the length of the retention.

Musical training and intensity discrimination: '11, R, .03, p.e., .03.

No correlation.

Musical environment and intensity discrimination: '11, R, .01, p.e., .03.

Musical enjoyment and intensity discrimination: '11, R, .01, p.e., .03.

Time sense and hearing ability: '16, r, .03, p.e., .04.

No significant correlation in these three cases.

Motor ability and hearing ability: '16, r, .24, p.e., .04.

This is a large correlation in view of the absence of specific common elements.

Free rhythm and hearing ability: '16, r, .01, p.e., .04.
Regulated rhythm and hearing ability: '16, r, .09, p.e., .04.
Rhythmic judgment and hearing ability: '16, r, -.06, p.e., .04.
No significant correlation in these two cases.
Singing keynote and hearing ability: '16, r, -.11, p.e., .04.
Singing intervals and hearing ability: '16, r, .32, p.e., .04.
Large correlation in the latter in view of the observed common elements.
Voluntary control and hearing abiltiy: '16, r, .00, p.e., .04.
Auditory imagery and hearing ability: '16, r, .03, p.e., .04.
No significant correlation in these two cases.
Motor imagery and hearing ability: '16, r, .22, p.e., .04.
No reason known for this correlation.
Tonal memory and hearing ability: '16, r, -.01, p.e., .04.
Motor ability and time-sense: '16, r, .01, p.e., .04.
No correlation.
Free rhythm and time-sense: '16, r, .08, p.e., .04.
Regulated rhythm and time-sense: '16, r, -.01, p.e., 04.
Rhythmic judgment and time-sense: '16, r, -.03, p.e., .04.
It is difficult to account for relative absence of correlation in these three cases.
Singing the keynote and time-sense: '16, r, .00, p.e., .04.
Singing intervals and time-sense: '16, r, .10, p.e., .04.
Voluntary control and time-sense: '16, r, .01, p.e., .04.
No noteworthy correlation.
Auditory imagery and time-sense: '16, r, .19, p.e., .04.
Motor imagery and time-sense: '16, r, .13, p.e., .04.
These correlations are significant because in time-sense both auditory imagery and motor imagery are conspicuous, according to introspections.
Tonal memory and time-sense: '16, r, .17, p.e., .03.
Tonal memory and time-sense are both memory tests.
Free rhythm and motor ability: '16, r, .09, p.e., .04.
Regulated rhythm and motor ability: '16, r, .12, p.e., .04.
The common element is general motor control.
Rhythmic judgment and motor ability: '16, r, .03, p.e., .04.
No correlation.

Singing the keynote and motor ability: '16, r, .10, p.e., .04.

Singing the keynote is, from one point of view, a motor control and it is safe to hazard the assumption that, to the extent that singing in key is at its maximum efficiency, the correlation between the motor ability and this capacity would increase. This point however remains to be established by further experiment; it represents a very important problem in vocational guidance as well as in pure psychology, or pedagogy.

Singing intervals and motor ability: '16, r, .21, p.e., .04.

The larger correlation here may indicate that the smaller correlation just noted for singing the keynote is exceptionally small. Theoretically the correlation for singing the keynote should be larger than this because it presents more of a problem of motor control.

Voluntary control and motor ability: '16, r, -.05, p.e., .04.

No correlation. Here again we are puzzled by the absence of correlation because precision in the control of the voice would seem to be related to motor ability in so far as it is independent of the hearing of pitch which is, of course, the largest determining factor.

Auditory imagery and motor ability: '16, r, .06, p.e., .04.

No marked correlation.

Motor imagery and motor ability: '16, r, .10, p.e., .04.

This correlation may be significant in that it may point to the existence of a general tendency which we may call motor-mindedness.

Tonal memory and motor ability: '16, r, .10, p.e., .04.

The explanation of this tendency towards correlation is uncertain.

Regulated rhythm and free rhythm: '11, r, .04, p.e., .05; '16, r, .14, p.e., .04.

It is difficult to account for the absence of correlation in '11, since these two factors involve several elements in common.

Rhythmic judgment and free rhythm: '16, r, .15, p.e., .03.

The relations are similar to those of regulated rhythm.

Singing the keynote and free rhythm: '16, r, .01, p.e., .04.

Singing intervals and free rhythm: '16, r, .07, p.e., .04.

Voluntary control and free rhythm: '16, r, .07, p.e., .04.

No significant correlation.

Singing and free rhythm: '11, r, .26, p.e., .05.

Correlation exceptionally large.

Auditory imagery and free rhythm: '11, r, .22, p.e., .05; '16, r, .01, p.e., .04.

It would seem probable that the correlation for '11 is more indicative of the true relation than the absence of correlation for '16.

Motor imagery and free rhythm: '11, r, .06, p.e., .05; '16, r, .07, p.e., .04.

Absence of correlation here seems strange.

Tonal memory and free rhythm: '16, r, .09, p.e., .04.

Musical environment and free rhythm: '11, R, .09, p.e., .03.

Correlation comparatively small.

Rhythmic judgment and regulated rhythm: '16, r, .00, p.e., .00.

Further refinements of these tests must be made.

Singing the keynote and regulated rhythm: '16, r, .13, p.e., .04.

Singing intervals and regulated rhythm: '16, r, .23, p.e., .04.

Voluntary control and regulated rhythm: '16, r, .11, p.e., .04.

Singing and regulated rhythm: '11, r, .07, p.e., .05.

The common element in these cases seems to be essentially the attitude of precision in action.

Auditory imagery and regulated rhythm: '11, r, .17, p.e., .05; '16, r, .00, p.e., .04.

Motor imagery and regulated rhythm: '11, r, .11, p.e., 05; '16, r, .08, p.e., .04.

The correlation for '11 in auditory imagery would seem to be a truer representative than the absence of correlation in '16; yet the low correlation in both auditory and motor imagery has great significance in the interpretation of the rôle of imagery in action.

Tonal memory and regulated rhythm: '16, r, .05, p.e., .04.

No significant correlation. From one point of view regulated rhythm involves memory; from another, it may be viewed as limited to the sensory span as distinguished from memory span.

Singing and regulated rhythm: '11, r, .07, p.e., .05.

Musical training and regulated rhythm: '11, R, .03, p.e., .03.

Musical environment and regulated rhythm: '11, R, .07, p.e., .03.

Musical enjoyment and regulated rhythm: '11, R, -.01, p.e., .03.

The absence of correlation in the last three cases raises a question about the significance of mechanically accurate appreciation of rhythm for musical enjoyment.

Singing the keynote and rhythmic judgment: '16, r, -.04. p.e., .04.

Singing intervals and rhythmic judgment: '16, r, -.05, p.e., .04.

Voluntary control and rhythmic judgment: '16, r, .08, p.e., .04.

No significant correlation.

Auditory imagery and rhythmic judgment: '16, r, .08, p.e., .04.

Motor imagery and rhythmic judgment: '16, r, .07, p.e., .04.

These low correlations would be reflections upon the efficacy of imagery as an aid to precision if we could regard the test of rhythmic judgment as reliable.

Tonal memory and rhythmic judgment: '16, r, .22, p.e., .04.

Memory is a basic factor in rhythmic judgment.

Singing intervals and singing the keynote: '09, R, .45, p.e., .04; '16, r, .61, p.e., .03.

Singing the scale and singing the keynote: '09, R, .42, p.e., .04.

Voluntary control and singing the keynote: '16, r, .55, p.e., .03.

Singing and singing the keynote: '09, r, .76, p.e., .02.

Singing the interval, singing the scale, and voluntary control involve the same factors as singing the key, plus the interval concept and memory for this concept. This correlation shows that, as a rule, those who can sing the key at all accurately tend to sing the interval accurately.

Auditory imagery and singing the keynote: '16, r, .24, p.e., .04.

Motor imagery and singing the keynote: '16, r, .19, p.e., .04.

These fairly good correlations would seem to indicate that those who are ear-minded or motor-minded have an advantage in that they sing in key.

Tonal memory and singing the keynote: '16, r, .32, p.e., .04.

The general sense of pitch is a common element.

Musical training and singing the keynote: '09, R, .14, p.e., .05.

a. *Instrumental:* '09, R, .11, p.e., .05.

b. Vocal: '09, R, .26, p.e., .04.

This relatively small correlation in the first two items is indicative of the fact that the immediate reproduction of a tone is perhaps one of the most elemental forms of singing and speaking and therefore responds the least to training. The difference between the correlation in instrumental training and vocal training cannot be laid entirely to the value of training in voice. It is perhaps more due to selection; those who have a naturally good voice control are the ones who train.

Musical environment and singing the keynote: '09, R, .24, p.e., .04.

Training, heredity, and selection are perhaps responsible for the correlation.

Musical expression and singing the keynote: '09, R, .23, p.e., .04.

a. *Instrumental:* '09, R, .15, p.e., .05.

b. *Vocal:* '09, R, .24, p.e., .04.

Explanation same as for instrumental training and vocal training.

Musical enjoyment and singing the keynote: '09, R, .04, p.e., .05.

Absence of significant correlation here is indicative of the fact that few persons are able to detect their error in singing or in the singing of others, and enjoy music as it comes, not knowing what fine distinctions they fail to make.

Singing the scale and singing intervals: '09, R, .47, p.e., .03.

Practically the same factors are involved.

Voluntary control and singing intervals: '16, r, .32, p.e., .04.

The problems in these two factors are more unlike than it would at first seem. Interval depends upon memory concept whereas the least producible difference depends more largely upon pitch discrimination and motor control.

Singing and singing intervals: '09, R, .75, p.e., .02.

This large correlation is due to the fact that singing intervals was one of the measurements of which singing is the average.

Auditory imagery and singing intervals: '16, r, .40, p.e., .04.

This large correlation is very significant. Most observers upon close introspection find that the singing of intervals is not a blind leap but the auditory image is actually present in concrete form before singing.

Motor imagery and singing intervals: '16, r, .21, p.e., .04.

The ratio of this correlation to the preceding one represents approximately the rôle of motor and auditory imagery in the singing of intervals.

Tonal memory and singing intervals: '16, r, .30, p.e., .04.

Tonal memory was measured in terms of memory of intervals.

Musical training and singing intervals: '09, R, .23, p.e., .04.

a. *Instrumental:* '09, R, .39, p.e., .04.

b. *Vocal:* '09, R, .15, p.e., .05.

While vocal training furnishes good exercise in the learning of the interval, the interval is probably learned far more through listening to music than through the voluntary adjustment of intervals as in singing. In other words, we get our intervals largely from the piano and other common instruments in which they are most frequently heard.

Musical environment and singing intervals: '09, R, .18, p.e., .05.

Correlation may be due to training, heredity, and selection.

Musical expression and singing intervals: '09, R, .24, p.e., .04.

a. *Instrumental:* '09, R, .21, p.e., .04.

b. *Vocal:* '09, R, .28, p.e., .04.

Explanation same as for musical training. Vocal expression is of course essentially the singing of intervals. There is, therefore, both training and selection whereas in piano playing, *e.g.*, it is chiefly selection.

Musical enjoyment and singing intervals: '09, R, .05, p.e., .05.

No correlation.

Singing and singing the scale: '09, R, .78, p.e., .01.

Musical training and singing the scale: '09, R, .18, p.e., .05.

Musical environment and singing the scale: '09, R, .16, p.e., .05.

Musical expression and singing the scale: '09, R, .31, p.e., .04.

Musical enjoyment and singing the scale: '09, R, .06, p.e., .05.

In singing the scale, the same principle operates as in the singing of the intervals, above.

Auditory imagery and voluntary control: '16, r, .29, p.e., .04.

Introspection shows very clearly that the auditory image is used in this test, often with profound consciousness, because there is a demand for imaging something which has not been heard.

Motor imagery and voluntary control: '16, r, .08, p.e., .04.

It is surprising that this correlation is not larger because introspection shows the motor image to be very conspicuous.

Tonal memory and voluntary control: '16, r, .50, p.e., .03.

This accords with introspections.

Auditory imagery and singing: '11, r. .07, p.e., .05.

Motor imagery and singing: '11, r, -.07, p.e., .05.

These absences of correlation suggest the need of more specialized inquiry into the actual efficiency value of images.

Musical training and singing: '09, R, .18, p.e., .05; '11, R, .09, p.e., .03.

None of the persons examined had enjoyed continuous professional training in singing; the fact that one had had a few more lessons than another is insignificant in comparison with the large differences in native capacity.

Musical environment and singing: '09, R, .22, p.e., .04; '11, R, .05, p.e., .03.

Musical expression and singing: '09, R, .32, p.e., .04.

Musical enjoyment and singing: '09, R, .01, p.e., .05; '11, R, .06, p.e., .03.

Compare these with correlation in singing intervals and voluntary control.

Motor imagery and auditory imagery: '11, r, .52, p.e., .04, '16, r, .42, p.e., .04.

All evidence points strongly toward a good correlation between auditory and motor imagery, much stronger, for example, than between motor and visual.

Tonal memory and auditory imagery: '16, r, .40, p.e., .04.

Naturally a good correlation. Tests of memory depend upon

the ability to hold the group of tones in a single memory image for comparison.

Musical training and auditory imagery: '11, R, .21, p.e., .03.

Musical environment and auditory imagery: '11, R, .13, p.e., .03.

Musical enjoyment and auditory imagery: '11, R, .11, p.e., .03.

Auditory imagery responds to auditory environment of which musical training is a special form.

Tonal memory and motor imagery: '16, r, .16, p.e., .04.

Many observers remember in terms of motor processes to an extent which would lead us to expect a larger correlation.

Musical training and motor imagery: '11, R, .04, p.e., .03.

Musical environment and motor imagery: '11, R, -.03, p.e., .03.

Musical enjoyment and motor imagery: '11, R, .11, p.e., .03.

The correlation in the last item is small in view of the fact that the enjoyment of music is interpreted to a considerable extent in motor terms.

Musical environment and musical training: '07, R, .26, p.e., .05; '09, R, .30, p.e., .04; '11, R, .28, p.e., .03.

There is a tendency for those who live in a musical environment to take musical training.

Musical expression and musical training: '07, R, .51, p.e., .03; '09. R, .48, p.e., .03.

While the correlation is good, it shows that musical expression is in large part independent of musical training.

Musical enjoyment and musical training: '07, R, .22, p.e., .05; '09, R, .03, p.e., .05; '11, R, .10, p.e., .03.

If these two factors are regarded as standing in the relation of cause and effect, this correlation is low.

Musical enjoyment and musical environment: '07, R, .03, p.e., .05; '09, R, .15, p.e., .05; '11, R, .13, p.e., .03.

Musical enjoyment and musical expression: '07, R, .12, p.e., .05; '09, R, .16, p.e., .04.

Musical enjoyment and instrumental expression: '07, R, .13, p.e., .05; '09, R, .20, p.e., .04.

Does musical education make a person hypercritical? Does recognition of failure in training tend to lower the estimate of musical enjoyment?

SOME GENERAL CONSIDERATIONS AND CONCLUSIONS

Certain general impressions have been gathered in the many years of work in this field; some of those that have direct bearing on the present situation should be mentioned.

Evolution in testing is a process of simplifying. In the test of consonance, *e.g.*, the first test was made with tuning forks and resonators under very complicated conditions of control. An elaborate system of weighting was devised for various degrees of consonance. All possible intervals within an octave were employed. The test was necessarily long. The definition of consonance was complicated. All these and some other requirements have given away to simple conditions. The test is now made with the piano. Only eleven intervals are selected and employed. The judgments are not weighted. The definition and the conception of consonance are simplified. Each of these steps in the direction of simplifying, except the resorting to the use of the piano, is a step in the direction of increased efficiency. This tendency is characteristic throughout our experience; improvement takes place through progressive simplifying of technique, delineation of problem, and statement of results.

The evolution of a test is a process of isolating and identifying a specific process. This is illustrated in the case of rhythm. Our so-called free rhythm and regulated rhythm are very unsatisfactory measures of rhythm as it appears in music and speech. Rhythm like memory falls into a number of distinct and separable factors. When we get each of these isolated and under control we shall perhaps not have any single test on rhythm as such, certainly not any to designate rhythm as a whole, for rhythm must be split up into its cognitive, affective, and motor components. Thus discrimination for time and intensity are basic. Temporal and intensive precision in action correspond to these on the motor side; but, in addition, mental content, tonal intellect, uninhibited impulse, temperament, subjective rhythm, and musical feeling in various aspects must be taken

into account. Thus, as the testing of rhythm progresses, we shall tease out, from that bewildering tracery of mental pattern which we call rhythm, the component fibres; and, by testing each of these in isolation, we shall be able to learn a number of things in particular about rhythm, just as we are now coming into a position to say a number of things in particular about a larger conception of musical talent as a whole.

Evolution in testing is a progressive mastery of the means of commanding attention and effort under the most favorable conditions. The charge and a personal command of attention and effort in the person tested are most important factors in the test. The conditions must be so set as to take advantage of known principles of attention and effort in the grasping of the situation and responding to it and the charge, whether written or oral, must be not only emphatic but a matter of record as an element in the result; and unfortunately only persons with a certain gift of proficiency in handling others, trained to the intricacies of sources of error, and alert to incidental evidences which may appear, can be trusted to apply mental tests effectively. There is great danger of scrutinizing the physical and methodological aspects of the test and overlooking the important rôle of the personal charge which is necessary to bring into play the mental factor which is to be tested.

Evolution in testing is a progressive freeing from ambiguities. The children or adults who come in to be tested have few of the concepts of the experimenter. Learning how to put one's self in the place of the one to be tested is a power but slowly acquired. The blind complacency on this point of many of those who engage in mental testing is amazing. Progress comes through the slow education of the experimenter in this wonderful art of divining the point of view of the person to be tested, and technically stripping apparatus, method, and directions of sources of confusion. The instructions are not satisfactory merely because couched in words of a common language.

Evolution in testing is in the direction of making the test elemental whenever we are dealing with an innate capacity as distinguished from an acquired ability. This is characterized

by the fact that to the extent that the test is elemental there should be no appreciable improvement with practice, development with age, or variation in general intelligence. Some tests may be made completely elemental. In others this characteristic can only be regarded as a goal which we may strive to approach. Of course the three factors named do not necessarily vary together. A test may be elemental in one respect and, in the others, only approach this in varying degrees. Since we cannot make all of the tests elemental for all three factors, it is important to know for any particular test to what extent the limit indicated by the record is merely cognitive in the sense defined in an earlier report (15), that is, does not represent physiological limit but is inferior on account of ignorance, lack of appreciation, unfitness of the test, and many other factors. Each of the tests here reported has been, as a rule, one hour in length. In this time it is possible to get permanently satisfactory records on the majority in a group, but there will always be a considerable per cent. who cannot be reached in that time. This is of course a grave source of error and is probably the chief source of discrepancy in correlation.

In view of the above considerations we are cautious in attributing significance to the correlations here found. In the period covered the tests have all been in a process of evolution. The correlations are themselves in many instances a criterion of the degree of efficiency in the test. In other cases the absence of correlation is a mark of accuracy in the tests. At best these findings should be regarded as a preliminary survey from which we may take our departure in more intensive critique of each test and its ramifications. We have found this preliminary survey exceedingly suggestive for guidance into further work and it is hoped that it may serve others in this field and in related fields in the same way.

With such reservations and apologies we may now attempt to gather some of the more salient features of interpretation for each test partly on the basis of the quantitative showing of the correlation but in general as a statement of impressions gathered in the experience with the tests. It should not be

regarded as a full summary of what has been stated in detail.

Pitch discrimination—The test of this capacity may be regarded as elemental, that is, the physiological threshold (15) is independent of variation with practice, age, sex, or general intelligence. It must, however, be remembered that in a first rough test, as here used, the approximate physiological threshold is reached only in somewhat more than half of the cases.

This capacity is basic for all perceptions of pitch and for production of pitch in singing and playing, *e.g.*, the violin; also for consonance, auditory imagery, and to some extent for memory. The failure of close correlation with these in the present tests may be attributed to numerous factors among which are the following: the related test does not involve the least perceptible difference in pitch, *e.g.*, singing or playing in pitch when the error is due largely to lack of motor control, or consonance and tonal memory in which keen pitch perception is not essential for a moderate performance; the related test may not be effectively differentiated, *e.g.*, singing an interval in which the chief source of error may lie in an erroneous conception of the interval; in cases of poor records the threshold may be merely a cognitive one and the actual capacity may be used in the related test, *e.g.*, in singing the keynote.

For hearing ability and discrimination for intensity, the common element may lie in both the structural or pathological conditions of the ear; but in time-sense, the various forms of rhythm, and related tests it may often be traced to common central conditions or inadequacy of one of the tests.

The data in the questionnaire are of special interest on these points: musical training furnishes relatively little training in precision of pitch perception; musical environment makes but little difference except in so far as it is selective, because the experience of pitch differences is called into play far more frequently outside of music than in music; ability to perform and enjoy music in a moderate degree is not seriously limited by defective pitch discrimination, except in a comparatively small per cent. of cases; the report of enjoyment of music cannot have taken into account what the reporter has missed by

not hearing, and may therefore represent a wrong conception of real enjoyment, jut as the color blind person has a wrong notion of what a color really is.

The measurement of pitch discrimination is undoubtedly facilitated by the presence of good auditory imagery, especially in a single group test. General intelligence also favors the securing of a good record. But there is no evidence to show that it correlates with actual psycho-physic capacity in pitch discrimination. In other words, lack of ear-mindedness and lack of general intelligence tend to produce unreliable record but they do not seem to effect actual psycho-physic capacity as functioning in familiar perceptions of sound. On the problem of heredity we have no satisfactory data.

Consonance.—Consonance is the most direct test we have on the ability to hear tone-quality. A test is now being developed for the measurement of the sense of timbre by measuring the ability to detect variations in the overtones. That ability is however closely related to the ability to judge the quality of two-clangs as in consonance, which is now the most general test of sensory capacity for musical intellect.

In view of this basic position of consonance as a test of the "ability to hear music" correlation with the returns from the questionnaire are of special interest. There is no evidence to show that those who have a natural sense of consonance are selected for musical training or that musical training and musical environment enhance this capacity; nor does a person's actual ability to hear tone-qualities as here measured seem to enter into his evaluation of his enjoyment of music. Our first move in the face of these facts was to question the validity and meaning of the consonance test. We are convinced of its validity both as to the isolation of factors concerned and fitness of method. The burden of squaring musical procedure with these facts is therefore passed on to the musical profession.

One of the most gratifying findings about this test is the surprising fact that it is elemental in several respects. This capacity is early present in bright children who have a musical ear. It is fairly independent of the vagaries of feeling, and

measures a well defined type of sensory judgment which is developed by the every-day reactions to sound. It is, however, in some respects an intelligence test as no one can make a good record without the ability to sustain attention purposefully in a difficult act of comparison. Lack of ear-mindedness is also a serious deterrent in this test.

Physiologically consonance presupposes a certain degree of pitch discrimination. It involves the same sensory attitude as intensity discrimination though far more sustained. It is undoubtedly selective with reference to the various measures of rhythm in that the sense of rhythm, which is of motor origin, is naturally developed by those who have a keen ear for tone quality. Between time-sense and pitch of singing there is no close relationship; but, for quality of singing and playing we shall of course look for a strong correlation in future tests since the expression of five points of quality rests upon the hearing of them.

Intensity discrimination.—In normal hearing this is a clear-cut test of the intellectual capacity for accuracy in the observation of sound. It is the simplest test in the entire series and is elemental as to effect of training (16), sex, and age. It varies however with intelligence, of which it is a measure in one aspect. In this respect it is a test of the power of concentration of attention in its simplest form. This capacity may, therefore, be regarded as an index to general ability and reliability in sensory observation of sound. This measure therefore has two phases of definite significance in the evaluation of musical talent: it indicates intensity discrimination and it is an index to general ability and reliability in hearing of sound. The latter point is well sustained by the significant correlation our figures show for so many of the factors here measured in which there is no other basis of relationship than that of precision in the observation of sound.

Musically the capacity must be distinguished from a lively and imaginative play of intensity in the appreciation of music. Artistic musical hearing, like painting, rests ultimately, not on rigid sensory impressions, but upon a fertile imagination which

takes the sensory impression as a cue. One who hears musically may be fantastic and fastidious in the intensity differences he attributes to sound and yet be inferior in the actual observation of such differences when they really exist; but, when it comes to singing or playing, this artistic license ceases and the effective performer is limited by his capacity for precision in the hearing of the shades of stress which he aims to convey.

Hearing ability.—This involves two variables, the relative sensitiveness of the physical ear as an organ and, as in intensity discrimination, the general power of attention to auditory perceptions. For the latter it is not so valuable as intensity discrimination because that is only one of the two variables and the other variable is the more important. Indeed the record of intensity discrimination should always be taken into account in the interpretation of hearing-ability as only one who has a good record in intensity discrimination can be counted on to reveal the actual physiological threshold in hearing-ability. The absence of correlation with so many of the tests here shown is therefore a mark of reliability in this test.

Time-sense.—There are three basic sensory aspects of music: pitch, intensity, and time. All other sensory aspects of music, such as timbre, consonance, and rhythm are complexes, in which one or more of these basic factors are elements. Time-sense as here used is the basic measure for the various temporal processes in the perception and expression of music. It complies with the requirements of a basic test in that it is simple and direct and is the most rigidly controllable measure available. Failure to correlate with the three so-called rhythm tests may therefore be regarded as a reflection upon the reliability of these, because time-sense is undoubtedly a basic capacity for all of them. The positive correlation with pitch discrimination and intensity discrimination is evidence of the presence of that common element which intensity discrimnation measures.

Motor ability.—Motor capacity in music as well as in any other form of diversified action may be analyzed into five fundamental phases, namely, motor ability, precision in time, intensity and direction of movement, simple and complex reaction time, strength and endurance.

The conventional measurement of motor ability by means of the tapping with a key has a specious appearance of accuracy. On account of its basic position in motor control it is extremely desirable that this test should be analyzed and standardized more accurately before being used extensively in diagnosis. The test should be so divided into parts as to measure both speed and regularity. As here used the test involved only speed. But speed and regularity are the common elements which are involved in all tests of motor control. If satisfactorily isolated in this test, they may be regarded as furnishing a two-phased measure of general ability in motor control. Together they are a mark of the general tone or condition of the motor organism. Singing and playing in pitch, reaction-time, rhythmic action, and endurance, each involves these two factors.

The greatest problem in the development of the evaluation of this test now is to determine to what extent general motor capacity as thus measured can be interpreted as an index to the motor control of specific organs, such as the finger, the lips, or the vocal cords.

One of the greatest obstacles in the way of such an evaluation is the fact that we do not know the relation between voluntary use of a capacity and its automatic use. Our tests are all more or less voluntary, conscious reactions, whereas in artistic performances and, indeed, in all skill, both sensory and motor processes are more or less automatic.

Free rhythm.—The term rhythm as used in this and the next test should be discarded, and we should speak of time,—the time of free action, the time of regulated action, in order to recognize, first, the fact that on the motor side these are time tests, parallel to time-sense on the sensory side; second, to dispel the notion that these are primary tests of rhythm, which they are not.

This test is psychologically simple and can be made relatively elemental. It is easier to control than motor ability, yet we are not entirely satisfied with the test as here reported. Both the '11 and the '16 free rhythm and regulated rhythm tests were made by students, comparatively inexperienced in testing, and it

was observed that they often failed in the very important matter of charge and command of mental attitude in the test. This situation emphasizes the fact that we may have perfected technique of physical method, and yet fail for want of personal command of the mental attitude in the observer. The notion that mental tests can be made so automatic as to eliminate the personal equation of the experimenter is absurd. This directs attention to the most urgent need in mental testing today—the need of not only formulating a charge, but of administering the test by a person who has strength of personality to make it effective, and technical insight enough to guard against unforeseen sources of distraction and variation in effort.

Regulated rhythm.—Psychologically, regulated rhythm is more complex than free rhythm and, therefore, is not so valuable. The failure to correlate with free rhythm and rhythmic judgment reflects unfavorably upon this test, because free rhythm represents the ability to play in correct time alone, and regulated rhythm the ability to play synchronously with another whose time is perfect, and rhythmic judgment is merely free rhythm one stage more complicated. These three tests, therefore, call for thorough revision.

We must be extremely cautious of the interpretation of the so-called rhythm tests, and not regard them as tests of rhythm in the sense of motor tendencies in the grouping of sounds by time and stress. The strongest kind of rhythmic person may be very inaccurate both in time and stress so far as the perception and feeling of rhythm is concerned, but when he attempts to convey his feeling to others by means of time he is limited by his power of precision in action which this test should measure.

Rhythmic judgment.—This test must be improved or rejected.

Singing the keynote.—This test is basic for all pitch singing. It is simple and precise, and, therefore, furnishes good opportunity for the accurate determination of many of the psychophysical principles involved in singing.

Most children who have a good ear develop this command

very early so that the first imitations of musical pitch may be true, and may appear to be as automatic as if they had been inherited as such. True pitch comes to a child as early and as naturally as a vowel quality in speech, where the child has a natural ear for music and has sensitive motor control. There are, however, many and extraordinary exceptions to this rule. Occasionally a very bright person apparently cannot tell "tune" and much less sing in tune at all, and yet be educable in the singing of pitch. Such cases seem to be due to a sort of tone ignorance analagous to that obliviousness to odors which many of us display in the presence of good normal sensitiveness to smell.

This test correlates well with the other tests of singing, and may, therefore, be regarded as a general index to voice control of pitch. At its best, it rests upon pitch discrimination. It seems to be favored by auditory and motor imagery, and by tonal memory. The correlation with vocal training, musical environment, and vocal expression is, perhaps, in large part a matter of selection.

All the singing tests can of course be paralleled on an instrument such as the violin.

Singing intervals and singing the scale.—The test of singing intervals may, of course, be a test of any kind of interval, such as a natural scale, chromatic scale, melody, or any one or more isolated intervals, conventional or otherwise. It is a natural, accurate, and very satisfactory test. If it is to be used as a general measure of the control of the interval concept, the singing of the natural scale is perhaps the most satisfactory because the most familiar. The singing of intervals is highly favored by auditory imagery and tonal memory, also, to some extent, by motor imagery. It correlates well with musical training, musical environment, and musical expression. The latter two are probably both selective and effective in their relation to the singing of intervals. No relationship with musical enjoyment is revealed.

Voluntary control of pitch in singing.—This is the complement

on the motor side to pitch discrimination on the sensory. This type of discriminative action gives promise of developing into a rôle possibly as important on the motor side as discrimination is on the sensory. This minimal sharping or flatting in singing can be paralleled in playing with any instrument of adjustable pitch. It answers the demand for a test that shall measure specific capacity in terms quite free from ordinary habit, and equally new to all. In this respect it differs radically from the singing of intervals or the keynote. Methodologically, it presents certain difficulties in securing a single measure in that there is no fixed standard as to the certainty of the change, but it seems that this difficulty can be overcome. This action involves primarily pitch discrimination and a form of motor control; it is favored by auditory imagery, motor imagery, and tonal memory, and it has many common elements with other singing tests.

Auditory imagery.—This is the only capacity in our series which does not lend itself readily to objective measurement. The introspective grading is, however, favored by the advantage of having fixed limits for the series of grades, namely, no image for the lower limit, and as vivid as true perception for the upper.

There are good unworked possibilities for the study of the nature of the true character of the musical make-up of an individual by an intensive individual study of the content, vividness, completeness, persistence, and spontaneity of the auditory imagery, particularly with the aid of some of the more recent forms of association tests.

It still remains to find out what rôle the auditory image really plays in the life of a musician, or in what respect it is essential to the highest appreciation of music. Unquestionably it holds a vital position in music, but the laws of its functioning and non-functioning are exceedingly complex.

In addition to the correlations already noted in preceding sections, we note here a conspicuous correlation with motor imagery and tonal memory. Records not included here show that motor imagery correlates far more closely with auditory imagery than with visual imagery.

Tonal memory.—This test is basic for the various phases of memory. It is clear-cut, rigid, simple, and adequate. If we shall have a single test for musical memory, this one answers the purpose, but a single test for such a complex act as memory must be regarded as a makeshift, since the various elements of memory must be isolated and tested separately, especially if intensive work is to be done.

Musical training.—In considering the returns of the questionnaire we have been cautious throughout to limit the applications to the degree of certainty implied in the character of the information itself. In correlating these facts with concrete measures from the objective tests it is equally important to limit the conclusions guardedly to the factor of talent actually controlled and measured, and to safeguard the conclusions in proportion to the accuracy of the analysis made by the test. Such precaution is necessary, not only for the writer, but it must also be borne in mind by the reader in his interpretation of statements like the following.

As far as the results of the questionnaire and supplementary evidences are reliable, we face a discouraging situation in the relation of the possession of musical talent and the opportunity for musical training. Briefly, we find that there is no satisfactory tendency to select the musically talented for musical training; that musical training does not ordinarily modify musical capacity; and that specific ability which might be developed by special training is not ordinarily enhanced by the training in vogue.

Musical environment.—The most impressive feature of the evidence before us on this point is the fact that the fundamental powers in musical talent such as pitch discrimination, intensity discrimination, time sense, imagery, and to some extent consonance, singing, and tonal memory, are developed quite apart from association with music. Musical training and environment seem to be a specialized form of the use of these capacities which are continuously in action in the appreciation of, and response to, an environment of sounds.

Musical expression.—It cannot be said from our evidence that

it is the musically talented who play and sing for us, although there is some correlation in that direction.

Musical enjoyment.—The real absence of correlation between musical enjoyment, and musical talent and musical training throws new light on the validity of prevailing opinion in regard to the reality and quality of our musical enjoyment. It is conceivable that a person who could hear and feel nothing but the rhythm of a drum beat might get real enjoyment from that and rest undisturbed in the notion that he had enjoyed the full riches of music, because, like the color blind, his system seems complete and he knows not what lies beyond.

In the light of the facts we have reviewed, we may ask the same question for musical theory in general that is proposed in regard to rhythm in the latest book on that subject (8): "Is it any wonder that the student who dips, or, more boldly, dives into the inevitable chapter on rhythm to be found in current musical hand books, rhetorics, treatises about versification, etc., emerges mystified, when so little account is taken of individual differences in what is one of the most 'individually different' of human experiences?"

It is difficult to convey in writing the illuminating insight into, and concrete foundations for musical theory, art, and pedagogy that one obtains in the prolonged use of these tests. But the general fact can be made clear that the psychology of music must be founded and organized upon experimental analysis of musical talent, that the aesthetics of music must be built in a scientific spirit upon such experimental analysis of the operating psychological principles, and that the art of teaching and vocational guidance in music must be based upon established facts about the nature of talent, the nature of the learning processes, and the nature of the aesthetic principles involved.

REFERENCES

1. ANDERSON, D. A. The duration of tones, the time interval, the direction of sound, darkness and quiet, and the order of stimuli in pitch discrimination. *Univ. of Iowa Stud. in Psychol.*, VI, 1914, 150-156 (*Psychol. Monog. 69*).

2. DUNLAP, KNIGHT. A new laboratory pendulum. *Psychol. Rev.*, 1912, XIX, 240-241.
3. FULLERTON-CATTELL. On the perception of small differences. Publ. of the Univ. of Pennsylvania, *Philos. Series*, 1892, (No. 2).
4. GAW, E. A. A revision of the Malmberg consonance test. (In this Volume.)
5. MALMBERG, C. F. The perception of consonance and dissonance. (In this Volume.)
6. MILES, W. R. Accuracy of the voice in simple pitch singing. *Univ. of Iowa Stud. of Psychol.*, VI, 1914, 13-66 (*Psychol. Monog.* 69.)
7. PATRICK, G. T. W. The new psychological laboratory of the university of Iowa. *Univ. of Iowa Stud. in Psychol.* III, 1902, 140-144.
8. PATTERSON, W. M. *The rhythm of prose.* Columbia Univ. Press, 1916, p. ix.
9. Ross, F. B. The measurement of time-sense as an element in the sense of rhythm. *Univ. of Iowa Stud. in Psychol.*, VI, 1914, 166-172 (*Psychol. Monog. 69*).
10. SEASHORE, C. E. An audiometer. *Univ. of Iowa Stud. in Psychol.*, II, 1899, 158-163.
11. SEASHORE, C. E. *Elementary experiments in psychology.* Henry Holt & Co., 1909.
12. SEASHORE, C. E. The measurement of pitch discrimination: a preliminary report. *Psychol. Monog.* No. 53, 1910, XIII, 21-60.
13. SEASHORE, C. E. A method of measuring mental work: the psychergograph. *Univ. of Iowa Stud. in Psychol.*, III, 1902, 1-17.
14. SEASHORE, C. E. The tonoscope. *Univ. of Iowa Stud. in Psychol.*, VI, 1914, 1-12 (*Psychol. Monog.* 69).
15. SEASHORE AND TAN. The elemental character of intensity discrimination. (In this volume.)
16. SMITH, F. O. The effect of training in pitch discrimination. *Univ. of Iowa Stud. in Psychol.*, VI, 1914, 67-104 (*Psychol. Monog.* 69).
17. SODERGREN, P. J. *The measurement of the memory span.* (Not yet published).
18. STEWART, R. M. The effect of intensity and order on the apparent pitch of tones in the middle range. *Univ. of Iowa Stud. in Psychol.*, VI, 1914, 157-160 (*Psychol. Monog.* 69).

THE PERCEPTION OF CONSONANCE AND DISSONANCE

BY

CONSTANTINE FRITHIOF MALMBERG

CONTENTS

Historical
- Definition and experimental methods
- Ranking of intervals

Analysis and ranking of intervals
- Preliminary study of apparatus
- Criteria of consonance and dissonance
- Method of procedure
- Statement of results

Measurements of the perception of consonance as a musical talent
- Apparatus
- Definition and illustration of criteria
- The weighting of demerits
- The norm established
- Bearing on the order of ranking
- Evaluation of the method

Recommendations toward a Standard Test
Summary of conclusions
Bibliography and references

The object of this investigation is first to establish the ranking order of the musical intervals within the octave c′c″ with respect to the degree of consonance, and second, to standardize a measurement of the perception of consonance.*

The term consonance has been variously defined, and has been used to convey several meanings. While in general, it

*The author wishes to express his hearty appreciation of the valuable assistance of the members of the Department of Psychology, particularly to the other observers named herein who gave their valuable services so freely. He wishes especially to mention his co-workers in the laboratory, Drs. Thomas Vance and Walter R. Miles and Messrs. Luther Widen, Felix B. Ross and Fred C. Bruene, whose valuable cooperation is highly appreciated. To Professor Seashore the author is under special obligations for the preliminary outlining of the problem, cooperation in the experiments, and much assistance in the preparation of this article.

has been used with reference to the agreement of simultaneous tones, we find each investigator or theorizer emphasizing some one factor in this complex phenomenon to the exclusion of one or two other factors of similar importance. This diversity of definition results in a corresponding diversity in the ranking order of the common musical intervals as to their degree of consonance. The method in this investigation will, therefore, be: first to determine the factors which enter into the perception of consonance, second, with these factors as a basis, to rank the intervals according to their relative degree of consonance or dissonance, third, to evaluate the ability to perceive consonance in terms of this ranking.

HISTORICAL

A brief resumé of the essential facts bearing most directly on this problem may be presented from the points of view of the theories and definitions of consonance, experimental methods, and ranking of intervals.

Definitions and experimental methods

Pythagoras discovered the regularity of the aliquot division of the vibrating string, and thereby gave a numerical value to the notes of the scale. Since then, the regularity of the vibration frequencies of tones has been known and accepted as a basis of consonance. No thought of the manner in which the mind perceives and distinguishes between consonances and dissonances occurred to the ancient Greeks; the perception was taken for granted. The pleasantness of some intervals and the unpleasantness of others were the only criteria that were present to their minds.

This conception prevailed even up to the time of *Leibnitz* (10) who was the first to call attention of scientists to the fact that the mind did not really analyze or perceive the actual number or the numerical regularity of the vibration frequencies in the intervals. Nevertheless, Leibnitz found no other explanation of this complex phenomenon in consciousness. Therefore, he appealed to the subconscious mind for a solution of the problem

and came to the conclusion that the mind unconsciously calculated the ratios of the vibration frequencies.

"Die Musik ist ein verborgenes Rechnen des Geistes, welcher nicht weiss, das es zählt. Denn er tut Vieles mit verwirrten oder unmerklichen Perzeptionen, was er in deutlicher Apperzeption nicht wahrnehmen kann. Die irren, welche meinen, es geschehe nichts in der Seele, dessen sie selbst nicht bewusst sei. Obwohl also die Seele nicht fühlt, dass sie zählt, fühlt sie doch das Ergebnis dieser unmerklichen Zählung, d. h. das aus ihr fliessende Vergnügen bei den Konsonanzen, Missvergnügen bei den Dissonanzen. Denn aus vielen unmerklichen Übereinstimmungen entsteht das Vergnügen."

This unconscious calculation produces a consonant interval when the ratio number does not exceed five. "Wir zählen in der Musik nicht über fünf", says Leibnitz.

Euler (2) agreeing essentially with Leibnitz' explanation, interpreted the feeling of agreeableness of the consonances as due to the ease of perceiving order or coherence in the simpler ratios. He divided the consonances into ten classes, ranking them according to the simplicity of their ratios. Euler was the first scientist to formulate the fundamental law of consonance that "the degree of consonance is in a direct ratio to the magnitude of the common divisor of the vibration frequencies."

Schopenhauer (24) explained the mental processes that accompany the perception of consonances and dissonances in greater detail and regarded music as the highest expression of the divine in the world. It owes its great power to the essential relation which it bears to the human will. Consonances are the result of a rational relation of the vibration frequencies of two notes that can be expressed by small numbers. Their constantly recurring coincidences can be apprehended more readily by consciousness than those whose coincidences are less frequent. The notes, which are the result of this relation, blend. Dissonance is just the reverse of this state, but both consonance and dissonance comprise different degrees, the one shading into the other.

"Music is a means of making rational and irrational relations of numbers comprehensible, not like arithmetic by the help of the concept, but by bringing them to a knowledge which is perfectly, directly and simultaneously sensible. Consonances and dissonances, with their innumerable degrees of difference, portray the movements of the human will in its essential feelings of satisfaction and dissatisfaction."

The distinguishing factor of consonance would accordingly be this feeling of satisfaction portrayed by objectifying the movements of the human will through harmony.

The method of investigation with regard to musical intervals was theoretical, supplemented by but meagre and indeterminate empirical data, up to the time of *Helmholtz* (4), who gave an impetus to experimental investigation through his analysis of sound in the attempt to demonstrate that consonance is dependent upon the coincidences of upper partial tones which result in a relative absence of beats. The important criterion of consonance for Helmholtz is *smoothness*. His definition reads as follows:

"Consonance is a continuous, dissonance, an intermittent sensation of tone. Two consonant tones flow on smoothly, side by side in an undisturbed stream; dissonant tones cut one another up into separate pulses of tone." Consonance is dependent upon "certain determinate ratios between pitch numbers which do not give rise to beats, or only such beats as possess so minimal an intensity as to produce no unpleasant disturbance of the united sound."

On this theory Helmholtz constructed his consonance curve, showing the relative degrees of consonance on the basis of the number of beats possessed by the different intervals as quoted in Table I below.

As early as 1751 *Tartini* (27) called attention to the importance of combination tones, but it was left for Preyer to prepare the way for Krüger's exposition on the influence of difference tones on consonance.

Preyer's (21) experiments with tones, whose overtones and combination tones were excluded, tend to establish the fact that

the feeling of pleasantness and unpleasantness of consonances and dissonances is dependent on the overtones and difference tones of the clang.

"The beats of the overtones and combination tones are," says Preyer, "a further criterion, for through them the smoothness of the sensation is destroyed. Yet they do not suffice to explain dissonances, as these occur without beats. The well recognized consonances must give the least combination tones, and the most unpleasant dissonances, the most; the former, the most coincidences, the latter, the least."

Stumpf is the pioneer in the purely psychological field, as he introduces a new point of departure. Stumpf (30) identifies consonance with tonal fusion:

"The sounding together of two tones approaches sometimes more, sometimes less, the impression of unity, and it is apparent that this is more the case, the more consonant the interval is. Even if we recognize the tones as two and separate from one another, yet they form a totality in the sensation, and this totality appears to us as possessing a greater or less degree of unity."

Thus Stumpf postulates "Verschmelzung" as the distinguishing criterion of the degree of consonance. He admits that other criteria exist, but this one factor is the only one necessary in ranking consonances and dissonances. Concerning this factor Stumpf writes as follows:

"Kann der Unterschied konsonanter und dissonanter Töne weder in unbewussten Funktionen noch in den Gefühlen liegen, so wird man ihn in den Tonempfindungen als solchen zu suchen haben, wo ihn denn auch Helmholtz suchte. Da er nun aber nicht in der begleitenden Obertönen und nicht in den Schwebungen liegen kann, so muss er eben in den beiden Tönen selbst liegen, welche wir konsonant oder dissonant nennen. Es ist, soviel ich sehe, nur ein Merkmal, das sich hier arbeitet: die *Verschmelzung* gleichzeitiger Töne."

As a correlate to his psychological theory, Stumpf proposes a physiological basis in his theory of "specific energy" which gives to each fusion its individual character.

Following out the theory that the better the consonance, the better is the fusion, Stumpf applied the technique of experimental psychology. That most consonant intervals are the most difficult to analyze into their constituent elements, is the basis of his method of analysis; and by this method he attempted to rank the intervals. He tested the perception of consonance with Appunn's Tonmesser and with the tones of the pipe organ, varying the quality of tone further by using the different stops of the organ, the observers being requested to record whether they perceived one or two tones, *i.e.*, whether or not the two objective tones served as one subjective. The number of errors for each two-clang constitute a measurement of the degree of consonance of that interval, and determined the ranking order.

Stumpf (29) found no difference between major thirds and sixths, yet he admits that there may be a fine degree of gradation between major and minor thirds and major and minor sixths. All the dissonances are ranked in one group as possessing the same degrees of dissonance, with the observation that the natural seventh, 4:7 may be a slightly better fusion than the other dissonances.

Stumpf (27) lays down the following laws of tonal fusion:

1. Fusion depends on the so-called ratio of vibrations.

2. The degrees of fusion are independent of the tonal region within the tonal range.

3. The degree of fusion is independent of the intensity, whether indeed it be the absolute or relative intensity, so long as the tones remain distinguishable.

4. The degree of fusion is not influenced by the addition of a third or fourth tone.

5. Very minimal deviations of the number of vibrations from the ratio create no perceptible difference in the degree of fusion. If the deviation is increased, the fusion in all pairs of tones, except the lowest degrees, passes into this degree without running through the intermediate degrees, if any. The rapidity of this transition is proportionate to the degree of the initial fusion.

6. Fusion remains and retains its degree when both tones do not affect the same ear.

7. Fusion remains in the mere representation of the imagination.

8. If we proceed above the octave, the same degrees of fusion recur with the ratio of vibrations increased one or more octaves. m : n 2x as m : n if m is less than n and x equals a small whole number.

Faist (3), following out Stumpf's method literally, attempted to verify these laws of fusion, in experiments on the pipe organ with the use of its different stops. In a preliminary experiment he employed what he termed "the direct method", *i.e.*, "the method of serial rank, in distinction from the indirect method used by Stumpf. This method involves the ranking of the intervals directly, keeping the whole series in mind and giving them their relative rank. The result of Faist's experiments tended to verify all of the laws postulated by Stumpf, except the third. He found that the relative intensity of the components of the interval did influence the perception of fusion.

Both *Meyer* (14) and *Stumpf* (34) have also attempted to measure consonance by means of reaction time. The consonances were always distinguished from the dissonances but this method showed no consistency in the ranking of the other intervals, and it was rejected by Meyer as unreliable. Meyer also investigated the effect of variations in intensity, and found that the less consonant the intervals are, the greater is the difficulty of recognizing them in their minimal intensity, and that the relative loss of intensity of higher tones synchronous with lower tones has no perceptible influence within the octave. As to the effect of presenting one tone to each ear, for the purpose of excluding the difference tones, Meyer affirms that this exclusion results in a loss of ability to rank the more difficult intervals, although consonances are readily distinguished from dissonances.

Buch (1) criticizes Stumpf for neglecting to regulate the conditions of his experiments. He thinks it unlikely that experiments made with the organ of the "Domkirche in Halle" could be carefully enough regulated to be trustworthy. Accordingly,

Buch constructed a special instrument, using the organ, by means of which he regulated the intensity, duration, and pitch of the intervals. His experiments were made with, and without analysis. He makes a distinction between making a judgment from the point of view of analysis and that of synthesis. He developed a twofold ranking of the intervals, one on the basis of fusion, and the other on the basis of smoothness.

Lipps' (11) explanation of consonance is based on the rhythmic coincidences of the tonal series. The rhythmic coincidences have their correlate in the psychic processes. For our immediate consciousness, consonance appears as an agreement in unity, or, as Lipps expresses it, "eine Zusammengehörigkeit, eine einheitlichkeit". Furthermore, it is an agreement that gives rise to a feeling of satisfaction. "Consonanz ist ein Verhältnis zwischen Tönen in dessen Natur es liegt, Befriedigung zu erzeugen." The most perfect consonance does not give rise to the most satisfaction as it is "empty and monotonous."

Meinong and Witasek (8) introduced the method of "paired comparisons" in determining the ranking of the intervals, as played on the tones of the violin. They state their conclusion briefly as follows:

"... zwei Töne um so mehr verschmelzen, (a) je näher ihnen der Klang steht, auf den als Partialtöne bezogen werden können, (b) je grösserer Zahlenwert dem Verhältnis ihrer Schwingungzahlen zukommt."

Their results are interpreted in terms of the Ebbinghaus theory of hearing. Their experiments demonstrate that the method of paired comparisons may be used to good advantage in the testing of the perception of the degree of consonance and the ranking of the intervals. Meyer has questioned the reliability of their results because of the inaccuracy of the violin as to pitch.

Using the method of "paired comparisons" with the tones of Appunn's Tonmesser, tuned to the accuracy of 1 v.d., *Pear* (19) determined the ranking of the intervals in degree of consonance. The observers recorded the intervals compared as equal, plus,

minus, or doubtful in preference, and the ranking order was computed on the basis of the number of "votes" given each interval. He considered fusion, analyzability, pleasantness and unpleasantness, and association, the factors which might enter into the perception of consonance. These factors were explained and illustrated for the observer, and he was instructed to make his judgment on the basis of fusion."

Wundt's explanation (41) shows that consonance is dependent on the confluence of various factors. There are four criteria that are more or less essential conditions of this phenomenon, namely: (1) Purity, the number of primary difference-tones of different orders, which combine to give the consonant chords a distinct or individual character; (2) uniformity, the uniform relation of the intervals to the compass of the scale; (3) the discrimination of consonance by the recognition of the tonal elements, dependent on the direct and indirect relation of clangs; (4) the fusion of tones into a "clang unity" through the dominance of one of the tonal elements—the one which arouses the most intensive associations.

Wundt explains dissonance as a "diffuse tonal fusion." The diffuse nature of dissonance arises, on the one hand, from the physiological condition of tone absorption; on the other, psychologically, through the distinct differentiation of tones arising from the compounding of the interfering difference-tones. Consonance is, therefore, an act of the apperceptive faculty of mind, which synthesizes the tones into a unity. The attention concentrates on the tonal element that carries with it the strongest associations, and brings all the related phenomena to a focus with the same.

Krueger (8) adopts the first criterion of Wundt as the explanation of consonance and dissonance. A two-clang always possesses five difference-tones. The pitch of each is determined by the formula h-l, 2h-l, 3h-l, etc. These difference-tones are related to every other simultaneous tone in the clang exactly in the same manner as the simultaneous primary tones. Thus, coincidences, beats, intermediate tones result from these difference-tones, and influence our perception of consonance. The

best consonances are those that are characterized by the absence of distinct difference-tones. Dissonances are the result of difference-tones in the two-clang which interfere with one another. The degree of dissonance depends on the number of these interfering difference-tones.

Krueger (6) bases his theory on empirical data. His method of experimentation had for its object the determination of the number and pitch of the difference-tones present in the various intervals. In order to exclude the overtones, he used the tuning forks. The observer was requested to record the number of difference-tones in each interval, and to identify the pitch of each on a Tonmesser.

Stumpf (29) has criticized Krueger's results as not showing sufficient consistency to substantiate his theory. As another proof against the theory he states that consonance and dissonance are perceived even when the tones composing the two-clang are presented one to each ear, thereby excluding the possibility of difference-tones.

To sum up the historical review, we have gathered the data into a table by different authorities and methods, showing the order of consonances and dissonances:

This historical summary shows that the factors emphasized in the course of investigating this complex phenomenon have been: the feeling of satisfaction, agreement of tones, smoothness, fusion, and purity, with slight variants of these. We recognize the fact throughout that the order of the intervals tends to correspond to the simplicity of the ratios, expressing this mathematical relationship. That the feeling of satisfaction, or the feeling of pleasantness, is too variable and general a factor to be used as a constant criterion, is evidenced by the fact that it has not entered as a determining factor in any of the experimental investigations.

The most fundamental factor in ranking consonance and dissonance may be termed *blending,* the tendency of tones to merge into a composite tone that shows a more or less distinct agreement of constituent parts in so far as they are perceived

TABLE I—*Historical order of consonances and dissonances*

Authors	*Date*	*Instrument*	*Method*	*Criteria*	No. Intervals.....	1	2	3	4	5	6	7	8	9	10	11	12
					Math. ratios*.....	c′ c″	c′ g′	c′ f′	c′ e′	c′ a′	c′ e′^{b}	c′ a′^{b}	c′ g′^{b}	c′ b′^{b}	c′ d′	c′ b′	c′ d′^{b}
Franco of Cologne	12th cent.		Theoretical	Pleasantness		1	2	2	3	4	3	4	5	5	5	5	5
Euler	1739		Mathematical	Pleasantness		1	2	3	4	5	6	7	8	9	10	11	12
Helmholtz	1863	Various	Coincidences of upper partials Beats	Smoothness		1	2	3	5	4	6	7	8	9	8	10	11
Stumpf	1883	Pipe organ (Different stops)	Analysis	Fusion		1	2	3	5	4	6	7	8	9	8	10	11
Faist	1897	Pipe organ (Different stops)	1) Direct 2) Analysis	Fusion Fusion		1	2	3	4	4	4	4	5	5	5	5	5
Meinong and Witasek	1897	Violin	Paired comparisons	Fusion		1	2	3	5	6	9	4	7	8	10		
Lipps	1899		Theoretical	Rhythmic coincidences		1	2	3	4	5	6	7	8	9	10	11	12
Buch	1900	Pipe organ Tonmesser	Analysis	Fusion Smoothness		1	2	3	3	3	4	4	4		5	6	
Krueger	1903	Tuning forks	Analysis	Purity		1	2	3	4	3	6	5	5	4	7		
Pear	1911	Tonmesser	Paired comparisons	Fusion		1	2	3	4	5	7	6		8	9	10	11

* The mathematical ratios are in some cases only approximate, *e. g.* the ratios for c′g′^{b} and c′b′^{b}, 5:7 and 9:5. The ratios differing slightly for the different authorities.

as members. This term has been used by some authorities as synonymous with consonance and to express the agreement of tones. The recurrent similarities of which the early scientists spoke and which Schopenhauer and Lipps have emphasized might also be classed under the category of blending. Preyer, Wundt, and Krueger have emphasized the importance of purity as a criterion and it has proved to be a specific mark of the agreement of the component tones.

Authorities agree with respect to the ranking of the octave and the fifth, first, and second, respectively; but, for the remaining intervals, we find disagreements due to the quality of tone, to the variation of the method of investigation, and to the variation in the basis of judgment.

It may never become possible to arrive at absolute agreement in the order of ranking, but it is plain from this brief historical survey that much may be gained in that direction by a clearer conception in regard to the nature of consonance, the analysis of conditions, and specific definition of terms for the purpose of experimental control. This will be the object of our next part.

ANALYSIS AND RANKING OF INTERVALS

The following series of experiments are the result of two years of investigation in the psychological laboratory of the State University of Iowa, extending from the fall of 1911 to the fall of 1913.

Preliminary Study of Apparatus

In a preliminary series of experiments, the reed organ, the pipe organ, blown bottles, the dichord, tuning forks, and the piano were tested with reference to their adaptability and efficiency in producing tones for an accurate measurement of the perception of consonance. Ten experienced observers were given the test, here designated as Series A, and were requested to give introspections with regard to the quality of tone. The recommendations which follow are based on this ranking and the introspections of the observers.

The reed organ was supplied with a set of accurately tuned

reeds, even temperament. The pipe organ pipes were of the open wooden type, producing a soft mellow tone. These were tuned to correspond to a set of forks tuned in just intonation, or forks of even temperament as the case required. The "blown bottles" were selected instead of the Stern tone-variators on account of their being more easily manipulated, *i.e.*, it was possible to have a complete series tuned permanently. This apparatus consisted of ordinary bottles of cylindrical shape, which, when empty, gave approximately the tone of c′, 256 v.d. They were tuned in just intonation by filling with paraffin. The mouthpieces were firmly attached to the bottles. The organ pipes and the "blown bottles" were energized by compressed air under constant pressure.

The dichord (Spearman's) was strung with a heavy piano string (wound wire), and was sounded by stroking with a cello bow. The wound string responds more readily, and the heavy cello bow eliminated the harsh and grating overtones and noises incident to the use of the ordinary wire string and the lighter bow of the violin. To avoid changes in the quality of the tone, the dichord was played in a uniform manner, care being taken to stroke the string evenly near the bridge.

In the early experiment, the forks were sounded mechanically before carefully attuned Helmholtz resonators, but later by a free movement of the hand. In the most successful mechanical devices, the handles of the forks were firmly mounted in rubber casings before the resonators, and were struck by means of hammers, mounted on steel springs. Wooden, cork, rubber and the regular felt piano hammers were tested, and the piano hammer was selected as producing the clearest tone. But, in spite of precautions, the thud incident to striking the fork, proved a distraction and caused impurities and, therefore, presentation by hand proved to be the best method.

The nine intervals, c′c″, c′e′, c′f′, c′g′, c′d′, c′e′, c′a′, c′b♭′, c′d′, were studied as sounded by each of these three instruments in two extensive series, (1) by the method of paired comparisons, and (2) by Stumpf's method of direct analysis; *i.e.*, in the former series, relative consonance was judged for successive

pairs upon each of the factors which may constitute consonance, and in the latter a direct judgment was based on fusion alone.

The reed organ was finally rejected because the tone was somewhat harsh and difficulty was experienced in securing a uniform timbre of the tone throughout the octave. The dichord was found impracticable as, beside showing variations in quality with slight variations of pressure and adjustment in bowing, it requires a change of the bridge for each note and is difficult to manipulate. The blown bottles, while producing a clear tone, presented difficulties in regulation. *Sylvester* (37) demonstrated that the pitch of a blown bottle is very difficult to control but we found that variations in the relative intensity of sound and the timbre were far more difficult to control. The inevitable difference in timbre made the tones stand apart in the two-clang in such a way as to be prohibitive.

The instruments available for the test were found to be the piano, the tuning forks, and the pipe organ.

The test of consonance with a rich quality of tone can be given more expeditiously by the piano than by the dichord. The piano offers an advantage in its availability and in the perfection it has reached as a musical instrument. It also has the advantage of being familiar and agreeable. Most of the sources of error, which may arise in the use of the piano, can be effectively guarded against. As three strings are sounded for each tone, each of these must be accurately tuned so that no beats arise in the single tone. The same end may be gained by damping two of the strings, thus leaving only one to vibrate. An accurate piano tuner can, however, eliminate the beats that are present in the single tones of the ordinary piano. To eliminate many of the impurities that may arise from resonance, the soft pedál should be used continuously in the test.

The organ pipes present the difficulty of regulating the wind pressure so as to keep it constant. As with the blown bottle, the variations in intensity and in the adjustment of lips causes considerable variations in the pitch. Uniformity of pressure and accurate adjustment is accomplished in the regular pipe organ by a rather complex mechanism, which it is difficult to

improve upon in the laboratory. It is, therefore, best to use a good organ where access to one can be had. The pipe organ of the Methodist Episcopal Church of Iowa City which we used. is a modern two-manual organ, which was in good condition. The most favorable stop for this experiment was found to be the "stopped diapason" which gives a tone intermediate in richness between the rich tones of the piano and the pure tone of the tuning fork. The tones of the organ are, however, not so clear-cut as the piano and tuning fork tones. Therefore, the observer has not the same certainty in his judgments of the organ tones as of the piano or tuning fork tones. Simultaneity in sounding the tones was regulated by opening and closing the stop after the keys of the interval had been pressed down.

The tuning forks are the most reliable in both pitch and timbre. If presented by hand, in a uniform manner, they give clear and distinct tones, which are especially well adapted for a test of consonance. The tuning forks used in these tests were accurately tuned to the tempered scale. The frequencies as recorded on the tonoscope registered: c′—258.6, d′♭—274, d′—290, e′♭ — 308, e′ — 326, f′ — 246, g′♭ — 364, g′ — 387, a′♭ — 410.14, a′ — 435.9, b′♭ — 460.8. b′ — 487.8, c″ — 517.2. Precautions were taken to secure pure tones by sounding the forks in a uniform manner before tuned Helmholtz resonators. The twelve resonators were mounted on the rim of a wheel and the c′ set on the center so that by giving a turn to the wheel any one of these in the rim could be swung into position horizontally to the right of the c′.

In the preliminary experimentation on the quality of the tone, two general facts were fairly well demonstrated. First, it was shown that the ranking of consonance will vary slightly for different qualities of tone. This is expressed quantitatively in later experiments. Second, when by the use of two sets of tuning forks, the just intonation was compared with the tempered intonation, no difference in ranking of the intervals large enough to affect the order resulted from the difference in temperament. This conclusion agrees with Faist's (3) statement with regard to minimal variations in vibration frequencies.

Criteria of consonance and dissonance

It is clear from the historical survey, and it was demonstrated in the preliminary experiments, that the fundamental reason for the great divergence in the ranking by experts and the consequent disparagement of the ranking of consonance and dissonance has been due to the failure to take common ground in the definition of these terms. Our first step was, therefore, to put the various claims of factors involved to a test in a long and painstaking series of analytical tests in which the various possible factors of criteria were isolated and discussed critically under control. This preliminary inquiry resulted in the recognition of the following factors:

For consonance:

1. Blending—a seeming to belong together, to agree.
2. Smoothness—relative freedom from beats.
3. Fusion—a tendency to merge into a single tone, unanalyzable.
4. Purity—resultant analagous to pure tone. (See Wundt.)*

For dissonance:

1. Disagreement—incompatibility.
2. Roughness—harshness, unevenness or intermittence.
3. Disparateness—separateness or seeming to stand apart—analyzable, "twoness".
4. Richness—resultant analogous to rich tone.

In terms of these factors we may then define consonance as follows: *When the two tones of a two-clang tend to blend or fuse and produce a relatively smooth and pure resultant, they are said to be consonant.* Dissonance is the reciprocal of this. "Agreeableness" which has played an important rôle in the popular conception and in the theory is here conspicuous by its absence. The perception of consonance as above defined there-

* Restfulness—a feeling of completeness, finality or satisfaction, with its opposite disquietude—a feeling of incompleteness, needing to be resolved, was first adopted as a fifth criterion, but it soon developed that it must be dropped as it is a variable criterion directly due to progression and association, which must be excluded.

fore becomes a cognitive act of discrimination rather than a mere feeling of agreeableness.

Method of procedure

The above definition, with its analysis into blending, smoothness, fusion and purity displayed on a chart, was read and discussed, giving all of the observers a consistent understanding of the definition.

The observers were instructed to judge each "two-clang" as an aesthetic object by itself, without respect to the effect of progression, meaning, association, or mood, and to make their judgments on each of the above given criteria in turn in separate series of experiments for each of the four criteria.

The experiment was made "with knowledge", and was conducted on the plan of an informal seminar, allowing a discussion of each judgment, but each observer recording his own final judgment. The intervals were sounded simultaneously, with a duration of approximately two seconds, and were repeated as often as requested by any observer—sometimes as many as fifteen or twenty times.

In determining the constant factors that enter into the perception of consonance and the ranking of the musical intervals, one must rely on the introspection and judgment of experienced observers of different types. For this experiment, therefore, eight observers were carefully selected on the basis of their training and fitness for the work.*

Even with the definition of consonance and the control of procedure here adopted, there remain many points of doubt and individual differences in opinion among observers as to the order of consonances. Since it was necessary to arrive at one

* They were the following: Professor C. E. Seashore, Assistant Professor Mabel C. Williams, Professor Edward Schaub, and Professor Robert Fullerton, Dr. Alma D. Schaub, Dr. Thomas Vance, Mr. Hazelette and the writer. Professors Seashore, E. Schaub, and Williams and Fullerton were members of the faculty, the last named being the head of the Department of Vocal Music. Dr. Thomas F. Vance was an advanced student in Psychology; Dr. Alma Devries Schaub had taken her Ph.D. in Psychology; and Mr. Hazelette was a graduate student in Physics and an experienced flute player.

particular order which might be considered a norm, the observers adopted the plan of sitting together in the experiment, proceeding very slowly, and discussing all cases of doubt or difficulty, analyzing the situation, varying the conditions of stimulation, and refining observations. This proved a very great advantage since each of the eight observers, of different types of training, offered to one another criticisms and suggestions for points of view and in this way distinctions were developed and errors of observation were eradicated which might otherwise have passed unnoticed.

Statement of results

As a group of trained observers, we found no difficulty in four-fifths of the sixty-six cases of paired comparisons at the beginning, and, after discussion, the differences of opinion centered about a still smaller number of the cases. Since each observer recorded for himself, we secured eight individual sets of ranking, although all of these were materially modified by the enlightment which came through mutual criticism and the repetition of trials. The record showed for each observer the ranking of the intervals for piano, tuning-fork, and pipe organ, and on each of the four factors, blending, smoothness, fusion, and purity separately. This series we may designate as *Series B*. The results will be stated in Table III after the results of the final test, *Series C,* have been stated in Table II.

It was the original intention to accept the average of these records (Table III) as a norm, but the discussion and mutual criticism was so stimulating and interesting that all the observers agreed to sit again and continue by the same method until all should agree and a unanimous verdict could be handed in as in the case of a jury. This was done with the piano and tuning-forks separately in what is here designated as Experiment Series C. The pipe organ was left out to shorten the labor, in view of the fact that the observers were of one mind that the piano and the tuning-forks were the best instruments available for the rich tones and the pure tones respectively.

The final returns in this series (C) are condensed in Table II,

where the record is kept for piano and organ for each of the eleven intervals on blending, smoothness, fusion, and purity in terms of the number of times a given interval was preferred. Since there were twelve intervals, the one that was preferred to every other would have a record of 11, meaning that it was preferred to eleven other intervals, the next one would be preferred to ten other intervals, etc. The actual rank is, therefore, approximately the inverse order of these numbers as is indicated by the Roman numerals.

The intervals are given in this table in the order in which they rank for the piano in the average.

It is manifestly out of the question to combine the ranking on fusion with the ranking of the other three criteria since the former results in a peculiar classification of its own. The other three criteria, however, seem to work together, prove mutually supplementary, and result in a fairly similar order of ranking. (See Table II). These are, therefore, brought together through the average in the last column of Table II, (Fig. 8), which may be regarded as the goal of the experiments in this Part on the order of ranking.

The data of this table are illustrated in Figures 1-8, which are self-explanatory and will aid materially in the interpretation of the table.

Table II is the principal table of facts to be considered here, but Table III is inserted for two purposes: (1) to furnish the tentative data for the pipe organ tones, and (2) to show to what extent the order of rank in Series C deviates from that of Series B (Table II).

The rank for the organ tone stated in Table III would probably not have changed very much if this tone quality had been carried through Series C as was done with the piano and tuning-forks. The change of relative rank for the piano and the tuning-forks from Series B to Series C may be seen by a comparison of these two tables, those in Table III representing the average of independent judgments of the eight observers and Table II the unanimous verdict reached later by the same observers.

While there are normal variations in the figures which denote

		Blending		*Smoothness*		*Fusion*		*Purity*		*Total* { Blending / Smoothness / Purity }	
	P.	11	I	11	I	11	I	11	I	11.0	I
c′c″	F.	11	I	11	I	11	I	11	I	11.0	I
c′ g′	P.	10	II	9	III	8	V	9	II	9.3	II
	F.	10	II	9	III	6	III	10	II	9.7	II
c′a′	P.	6	VI	10	II	3	VIII	7	IV	7.7	III
	F.	7	VI	10	II	2	VII	8	III	8.3	III
c′ e′	P.	9	III	7	V	6	V	5	V	7.0	IV
	F.	8	IV	7	IV	7	IV	5	V	6.7	V
c′ f′	P.	7	V	6	VI	5	VI	8	III	7.0	V
	F.	9	III	7	IV	7	IV	5	V	7.0	IV
c′ a″^{b}	P.	5	VII	8	IV	3	VII	5	V	6.0	VI
	F.	6	VI	7	IV	3	VI	6	IV	6.3	VI
c′ e″^{b}	P.	8	IV	5	VII	8	IV	2	VIII	5.0	VII
	F.	5	VII	3	VII	7	IV	3	VI	3.7	VII
c′ g″^{b}	P.	4	VIII	4	VIII	4	VII	3	VII	3.7	VIII
	F.	4	VIII	5	V	4	V	3	VII	4.0	VIII
c′ b″^{b}	P.	3	IX	3	IX	1	IX	4	VI	3.3	IX
	F.	3	IX	4	VI	1	VIII	3	VII	3.3	IX
c′ d′	P.	2	X	1	XI	9	III	2	VIII	1.7	X
	F.	2	X	1	IX	8	III	1	IX	1.3	XI
c′ b′	P.	1	XI	2	X	0	X	1	IX	1.3	XI
	F.	1	XI	2	VIII	0	IX	2	VIII	1.7	X
c′ d″^{b}	P.	0	XII	0	XII	10	II	0	XI	0.0	XII
	F.	0	XII	0	X	10	II	0	X	0.0	XII

TABLE III—*Rank of consonances and dissonances (Series B)*

		Blending		*Smoothness*		*Fusion*		*Purity*		*Total* { Blending Smoothness Purity	
		av.	m. v.	av.	m. v.	av.	m. v.	av.	m. v.	av.	m. v.
	P.	10.9	.2	11.0	.0	10.5	.7	11.0	.0	10.9	.2
c′ c″	F.	11.0	.0	11.0	.0	11.0	.0	11.0	.0	11.0	.0
	O.	11.0	.0	11.0	.0	10.8	.3	11.0	.0	10.9	.1
c′ g′	P.	9.0	1.0	9.0	1.0	5.6	.9	9.3	.9	8.2	.9
	F.	9.9	.2	9.4	.5	6.3	1.6	9.6	.5	8.8	.7
	O.	9.6	.6	9.4	.8	6.6	.9	9.3	.6	8.7	.7
c′ a′	P.	6.8	.5	7.5	.8	3.0	.5	7.3	.6	6.2	.6
	F.	6.0	.7	7.1	.7	2.8	.8	7.5	1.0	5.9	.8
	O.	6.0	1.0	6.8	1.3	2.8	.6	7.0	1.0	5.7	.7
c′ e′	P.	8.8	.6	7.5	1.1	7.1	.4	7.3	.9	7.7	.8
	F.	8.4	.5	7.5	.6	7.0	1.0	7.0	1.0	7.5	.8
	O.	9.0	.3	7.1	1.2	6.5	1.3	8.0	.8	7.7	.9
c′ f′	P.	7.6	1.0	6.8	.8	4.9	.7	7.4	1.3	6.8	.9
	F.	7.9	.2	6.0	.8	6.0	.3	6.5	.8	6.6	.5
	O.	7.9	.5	7.3	1.1	5.1	.7	6.9	.9	6.8	.8
c′ a′♭	P.	6.1	1.2	6.5	1.0	3.1	.7	6.4	1.3	5.5	1.0
	F.	5.5	.8	6.5	.9	2.8	.4	6.6	.8	5.4	.7
	O.	4.6	.5	5.3	.9	3.1	.7	5.9	.4	4.7	.6
c′ e′♭	P.	6.1	.4	4.8	1.0	6.1	1.4	4.1	.7	5.4	.8
	F.	5.9	.2	2.5	.8	7.5	.6	4.3	.6	5.1	.6
	O.	6.9	.4	5.0	.8	7.0	1.0	4.4	1.0	5.8	.8
c′ g′♭	P.	4.0	.0	4.1	.2	3.5	.8	4.3	.9	3.9	.5
	F.	4.9	.7	4.4	.9	4.1	.7	3.8	.4	4.8	.7
	O.	4.8	.8	4.8	.6	4.1	.9	4.0	.4	4.5	.7
c′ b′♭	P.	2.6	.5	2.9	.2	1.0	.0	3.6	1.3	2.5	.5
	F.	3.0	.0	3.8	1.9	1.0	.0	3.8	1.2	2.9	.8
	O.	3.0	.3	3.1	.9	1.6	.6	3.6	1.7	3.8	1.1
c′ d′	P.	2.4	.7	1.4	.5	6.9	1.7	2.0	.3	3.2	.8
	F.	1.9	.2	1.0	.0	7.0	.5	1.0	.0	1.4	.2
	O.	1.3	.4	1.3	.4	8.1	.7	1.3	.4	3.0	.6
c′ b′	P.	.6	.3	1.5	.5	.3	.0	1.0	.0	.9	.1
	F.	1.1	.2	2.0	.3	.0	.0	2.1	.2	1.3	.2
	O.	2.0	.3	1.8	.6	.3	.2	2.8	1.5	1.7	.3
c′ d′♭	P.	.0	.0	.0	.0	9.8	.4	0.0	.0	2.5	.1
	F.	.0	.0	.1	.1	9.8	.4	.1	.1	2.5	.2
	O.	.0	.0	.0	.0	9.3	1.0	.0	.0	2.3	.3

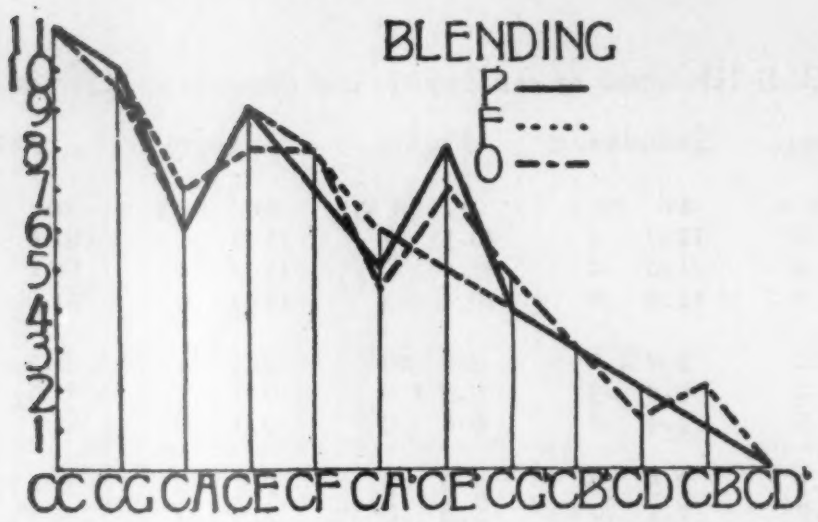

FIG. 1. Blending, for piano, forks, and organ respectively.

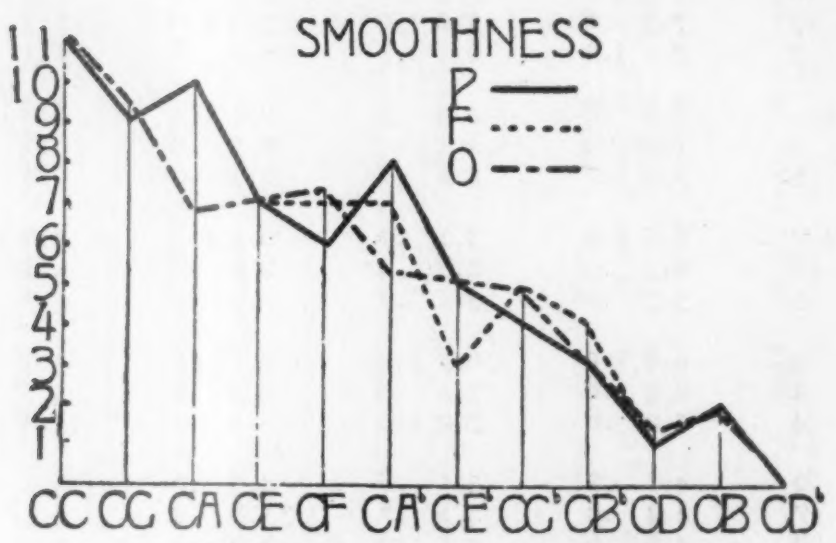

FIG. 2. Smoothness, for the three instruments.

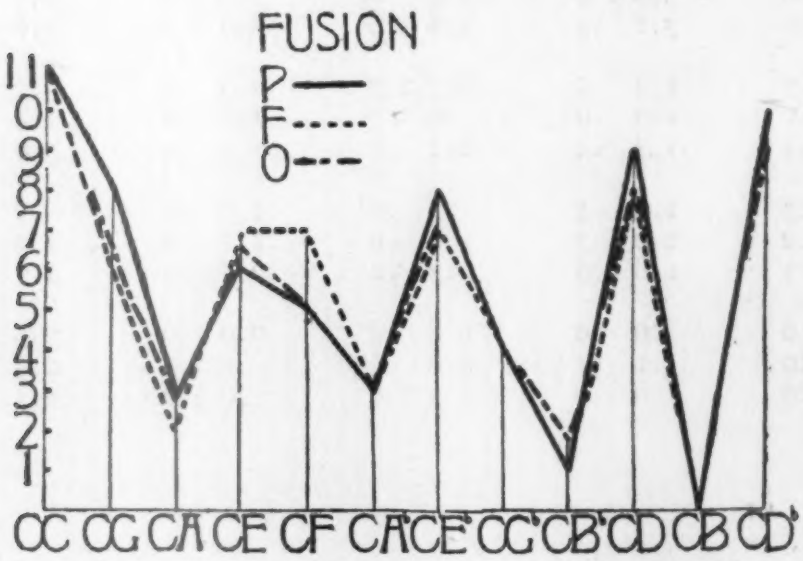

FIG. 3. Fusion, for the three instruments.

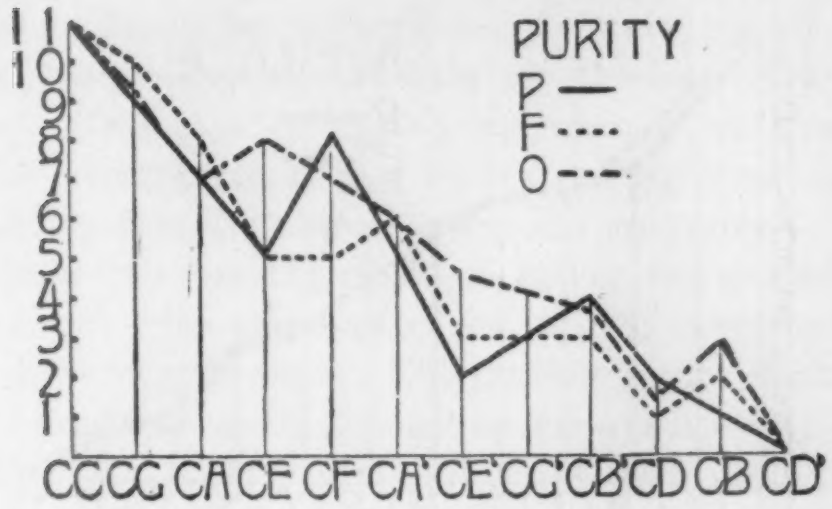

FIG. 4. Purity, for the three instruments.

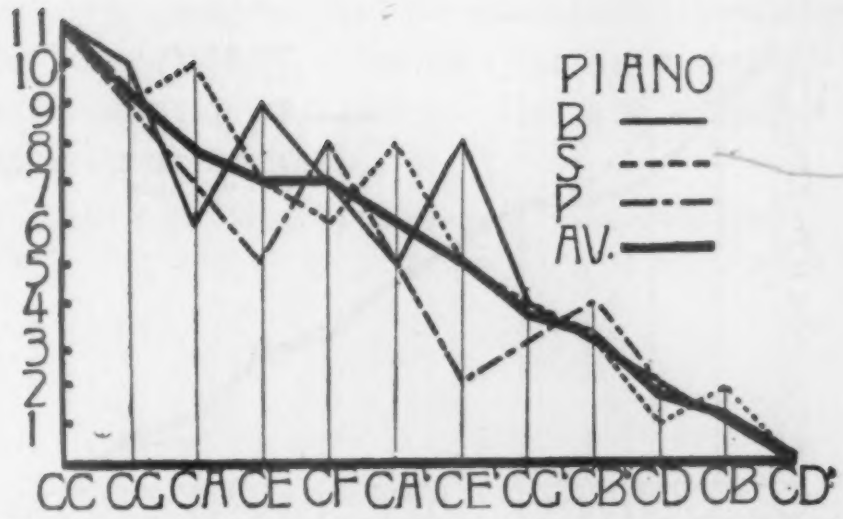

FIG. 5. Blending, smoothness, and purity, with the composite (Av.) from these for the piano.

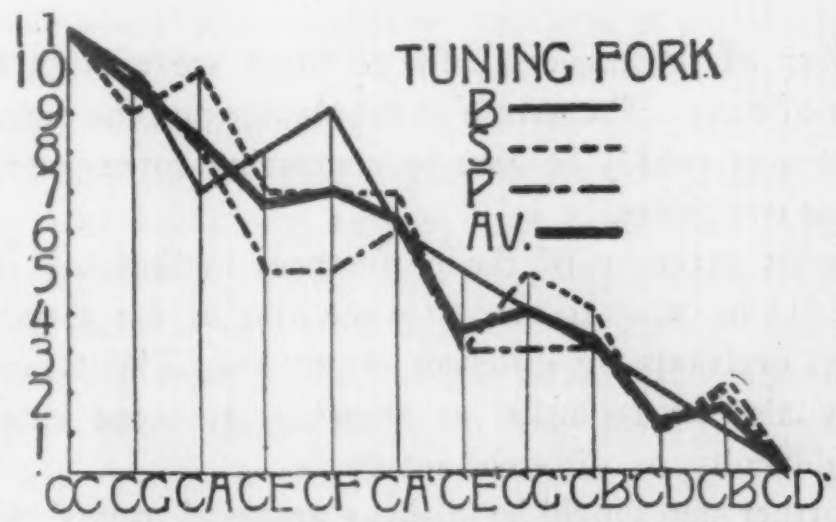

FIG. 6. Blending, smoothness, and purity, with the composite from these for the tuning forks.

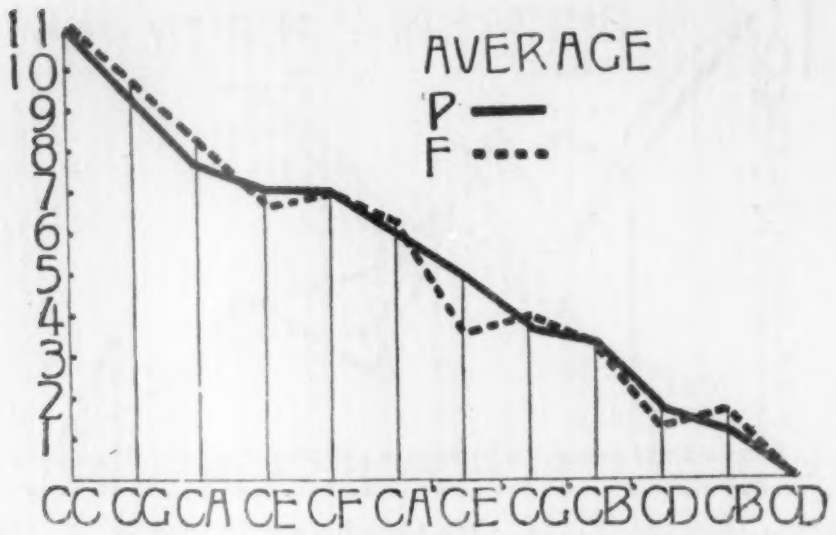

FIG. 7. Comparison curves for piano and tuning forks.

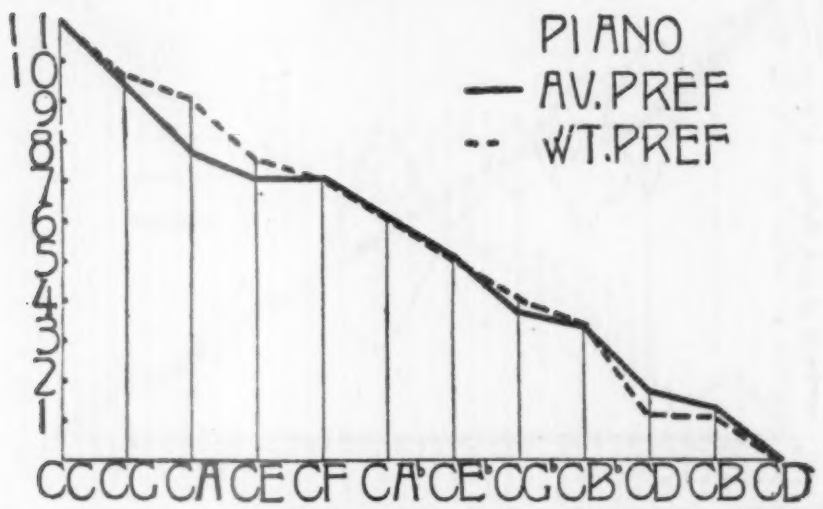

FIG. 8. Comparison of the average preference and the weighted average preference for the piano.

the number of preferences, few of these are enough to alter the order of rank. The nature of the change and the quantitative difference may readily be seen by comparing corresponding figures in the two tables.

The mean variation of the preferences in Series B is given in Table III in order to convey some idea of the variability in individual decisions by different observers. The mean variation may also be regarded as showing, to some extent, the relative difficulty of different intervals.

One further step should be taken: a group of trained observers should proceed as in Series B and C above, except that only

one decision should be rendered for the three criteria on the following basis: *Give the decision on blending alone if the degree of blending is perceptibly different; if not, make the decision on smoothness; and, if there is no difference in either smoothness or blending, base the decision on purity.*

That procedure has the merit of basing the decision on a single criterion—the appropriate one in each case—instead of striking an average of three. The ranking by this method may possibly modify the order denoted by the average in Table II.

To what extent there is likelihood of any effect upon the order of rank by this step may be judged upon analyses of the basis of the decision for each part of intervals as recorded in Series C. In this series the group worked for a unanimous verdict on what constituted the deciding criterion, just as it worked for a unanimous verdict on the order. The final decisions of the group are contained in Table IV.

TABLE IV—*Basis of judgment for individual pairs*

	c′ g′^{b}	c′ d′^{b}	c′ f′	c′ d′	c′ e′	c′ e′^{b}	c′ g′	c′ c″	c′ a′^{b}	c′ b′	c′ a′
c′ d′^{b}	s										
c′ f′	b	s									
c′ d′	s	s	s								
c′ e′	b	s	p	s							
c′ e′^{b}	b	s	p	s	p						
c′ g′	b	s	b	s	s	s					
c′ c″	b	s	b	s	p	p	b				
c′ a′^{b}	s	s	b	s	b	s	b	b			
c′ b′	b	s	b	b	b	b	b	b	b		
c′ a′	s	s	s	s	s	s	b	b	s	b	
c′ b′^{b}	b	s	b	s	b	b	b	b	s	b	s

b—blending; p—purity; s—smoothness. The order of the intervals in this table is the order used in giving the test.

Table IV shows that thirty-one of the pairs are determined primarily by the factor of blending, thirty-one primarily by the factor of smoothness, and only in four pairs does purity enter in as the determining factor. In the comparisons of the dissonant intervals c′d′ and c′d′^{b} with the remaining intervals the element of roughness is the characteristic that determines the judgment in every case, except in the pair c′b′ and c′d′. c′c″ and c′a′ depend primarily upon the criterion of blending. In the comparison with c′e′ and c′c″, purity enters as the deciding factor. In comparison of c′a′ with these two intervals,

smoothness gives c′a′ the first place. The ranking of c′a′ is largely determined by the criterion of smoothness. However, in the judgments of the pairs c′a′ and c′b′, and c′a′ and c′b′♭, blending becomes the criterion. c′e′ is determined by blending in its comparisons with c′b′♭, c′f′, c′a′♭, c′b′, c′b′♭. In the judgment of the difference of major and minor thirds, purity is the most distinct element which gives c′e′ the better rank. c′f′ in all cases not mentioned above, except in the case of c′e′♭, is judged on the blending factor. In the case of c′e′♭ the purity of the two intervals determine their rank, as c′f′ ranks II in purity while c′e′♭ ranks III. The minor sixth ranked above the minor third on account of its relatively greater degree of smoothness. In all other cases, not mentioned above, the minor third is judged on the basis of blending. c′b′ is judged on its lack of the factor of blending in all cases, except in its comparison with c′d′♭. In the two dissonances c′b′ c′d′, there is difficulty in deciding on account of their apparent difference in character. Both are extremely dissonant, c′b′ ranking lower in blending and c′d′, on the other hand, ranking perceptibly lower in smoothness. As the blending is the more important factor, c′d′ is given the first place. In the pair c′b′♭ and c′a′♭, there is a minimal difference in blending, but c′a′♭ is perceptibly smoother than c′g′♭. In the intervals c′g′♭ and c′b′♭ blending determines the judgment. In some cases, such as c′c″ and c′a′, all three criteria coöperate. The superior ability to perceive consonance consists in selecting the factor which most influences the ranking.

Turning then to the specific question before us, we may estimate, on the basis of data in Table IV, what the probable effect would be of a single decision according to the above directions as compared with an average decision for these three criteria.

The analysis is reduced to final figures in Table V in which "average preference" is taken from the last column in Table II for piano and tuning forks respectively;—B., S. and P. denote the number of times the judgment was based upon the factor of blending, smoothness, and purity respectively, for each interval when compared with other intervals according to the exhibit in Table IV.

"Weighted preference" is calculated by weighting the average in accordance with the number of decisions based on each of these factors. Thus, turning to Table IV, we see that for the interval c′g′ the decision may be reached seven times on blending, and four times on smoothness, while purity need not be considered; the preference rank for blending is 10 and for smoothness 9. For this interval, we accordingly get 7x10+4x9 ÷11=9.6 as the weighted average.

Section A, of table V (Fig. 8) is based on the piano quality but as a rough approximation we have used the same weighting for tuning forks, Section B of Table V.

TABLE V.—*Comparison of weighted and average preferences*

A. For Piano

Int.	Av. Pref.	Order	Wt. Pref.	Order	B.	S.	P.
c′ c″	11.0	I	11.0	I	6	2	2
c′ g′	9.3	II	9.6	II	7	4	
c′ a′	7.7	III	9.0	III	3	8	
c′ e′	7.0	IV	7.5	IV	5	4	2
c′ f′	7.0	V	6.9	V	7	3	1
c′ a′^{b}	6.0	VI	5.9	VI	5	6	
c′ e′^{b}	5.0	VII	5.0	VII	3	5	3
c′ g′^{b}	3.7	VIII	4.0	VIII	7	4	
c′ b′^{b}	3.3	IX	3.0	IX	7	4	
c′ d′	1.7	X	1.1	X	1	10	
c′ b′	1.3	XI	1.1	XI	10	1	
c′ d′^{b}	0.0	XII	0.0	XII		11	

B. For Tuning Fork

c′ c″	11.0	I	11.0	I
c′ g′	9.7	II	9.6	II
c′ a′	8.3	III	8.3	III
c′ e′	6.7	V	8.0	V
c′ f′	7.0	IV	8.1	IV
c′ a′^{b}	6.3	VI	6.5	VI
c′ e′^{b}	3.7	VIII	3.5	VIII
c′ g′^{b}	4.0	VII	4.4	VII
c′ b′^{b}	3.3	IX	3.4	IX
c′ d′	1.3	XI	1.1	X
c′ b′	1.7	X	1.1	XI
c′ d′^{b}	0.0	XII	0.0	XII

It is rather surprising that for the piano the weighting does not alter the order found for the mere average. For the purpose of a tentative working norm with the piano, tempered scale, (when the decision is made on one of the three factors directed as above, namely: if possible, on blending; if not, on smoothness; and if not on either of these two, on purity), the

order in which the intervals are given in Tables II, III, and V may be regarded as the standard order from the best consonance to the worst dissonance, as expressed in Fig. 9.

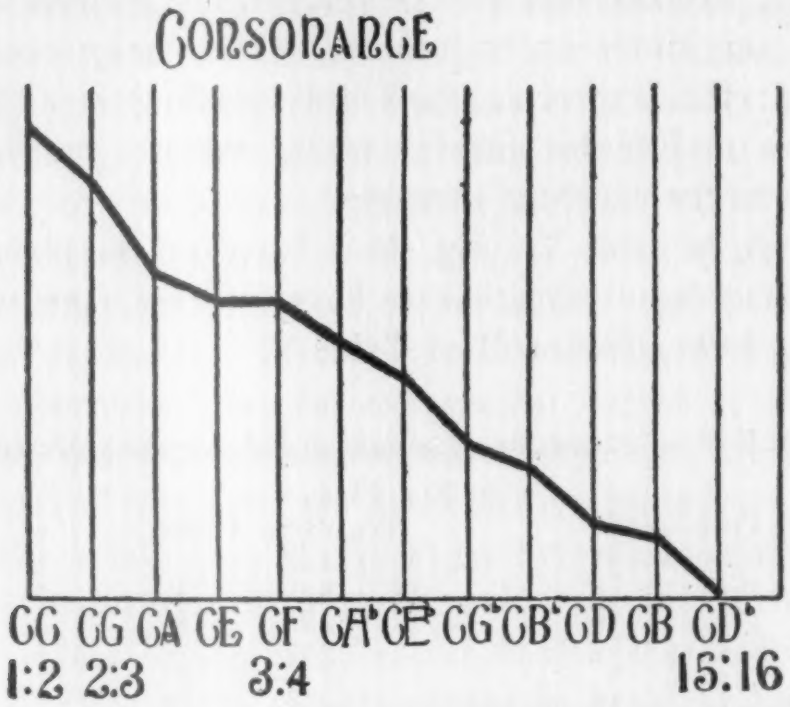

FIG. 9. Norm: the order of consonance-dissonance for the piano.

A rough approximation to the relative degree of certainty in the preference of one interval to another is shown by the numbers which denote the number of times an interval is preferred. This is true both for great differences, *e.g.*, c′c″ and c′d′, as well as for adjacent intervals, *e.g.*, c′c″ and c′g′. The irregularities in the curve, of course, mean just such differences; thus, in Fig. 9 the difference between the consonance of c′c″ and c′g′ is decidedly larger than the difference between c′e′ and c′f′.

Since the primary object of this investigation was to standardize the measurement of the sense of consonance, the main object of the present section of the work was to secure a norm which might be considered a conventional standard for such purposes. All this work on ranking is, therefore, merely accessory to the main object of this investigation. For this reason, the detailed discussion of the introspections and the theoretical interpretation of the empirical records may perhaps best be deferred for a more elaborate investigation.

We have here, at least, a tentative norm established for the first time after recognition of the principal factors isolated under experimental control. Later refinements of experiments may make minor changes in the order, but we have made progress by developing a working principle—the recognition of specific criteria.

MEASUREMENTS OF THE PERCEPTION OF CONSONANCE AS A MUSICAL TALENT

Having determined the constant factors of consonance, the ranking order of the intervals, and the method and apparatus adapted to this test, the writer will in this section, give an account of measurements made under the above prescribed conditions. These measurements were made with three ends in view: first, to secure measurements of individual abilities; second, to establish a norm; and third, to test this measurement under controlled conditions.

Apparatus

The intervals of the octave c′c″ were presented by the method of paired comparison, being played on a piano accurately tuned in the tempered scale. The tones were sounded with moderate loudness, the soft pedal being applied continuously during the test. Each two-clang was sounded with a duration of approximately 2 seconds, with an interval of 1 second between the members of each "pair", and 4 seconds between each repetition. Every pair was sounded at least twice, more frequently three or four times, and in the more difficult comparisons a relatively greater number of times.

Definition and illustration of criteria

Special precautions were taken to impress the fact that mere agreeableness is not the basis of the decision. It was pointed out that a dissonance may be very agreeable for some musical purposes and that the rich musical body of the two-clang, such as the major third, may be the most agreeable and still not the most consonant. It was made clear that the test is a cognitive act of discrimination as opposed to the traditional affective test employed in music.

A placard bearing the definition of consonance was placed before the class, reading as follows:

Consonance is:

(1) *Blending,* a seeming to agree, to belong together.

(2) *Smoothness,* relative freedom from beats.

(3) *Purity,* thinness of tone, absence of richness.

The conductor of the test then explained tersely and concretely how the experiment would proceed, how to record, and how to apply the above-defined criteria in arriving at a decision. The 66 trials, which constitute one complete set, were made in one hour. The intervals used to illustrate the above criteria were sounded in a different octave from the one used in the test. Both the negative and positive aspects of the criteria, *i.e.,* consonance and relative absence of consonance were presented. Thus, without naming the intervals, c′c″—c′b′ and c′e′—c′b′ were given to illustrate blending, the first interval in each pair bringing out the positive, the second, the negative aspect of blending. As illustrations of smoothness, the pairs c′b′—c′d′ and c′a′—c′e′^{b} were played. In both of these pairs judgment is best based on smoothness, the first presenting a large difference, the second a small difference. The element of roughness in the minor third is scarcely perceptible, but it is the deciding factor in the comparison of these two intervals, as the major sixth undoubtedly owes its higher rank to its greater degree of smoothness in this pair. c′c″—c′e′^{b} were selected to illustrate tonal purity. The octave is a relatively pure interval, while the minor third is rich in quality. In order to give a still clearer conception of purity, a criterion somewhat vague in the mind of the average observer, a tuning-fork was sounded before an attuned Helmholtz resonator, which gave a clear pure tone; and this was compared with the rich tone of the piano and violin and with two-clangs. The observers were finally instructed to make their judgments on each pair with respect to the three criteria in the order named and illustrated; namely, giving first preference to blending, second to smoothness, and third to purity. In all cases where they found blending to be the decisive factor, they were instructed to judge on this criterion. In case they

were unable to decide on the basis of blending, they should take the second criterion into account; and, if smoothness was not decisive, the test of purity should be applied. In a large portion of the cases the three criteria coöperated and agreed.

Since the sequence of intervals is a factor that cannot be overlooked, a fixed order of comparisons, which distributed the consonances and dissonances fairly was adopted, as given in Table VI. The body of this table contains three sets of information:

TABLE VI—*The order of trials, the correct preference and the schedule of demerits*

	c′ g′♭	c′ d′♭	c′ f′	c′ d′	c′ e′	c′ e′♭	c′ g′	c′ a′♭	c′ b′	c′ a′
	1									
c′d′♭	1 4									
	2	3								
c′ f′	2 3	2 7								
	22	4	5							
c′ d′	1 2	2 2	1 5							
	23	24	6	7						
c′ e′	2 4	2 8	2 1	2 6						
	39	25	26	8	9					
c′ e′♭	2 1	2 5	1 2	2 3	1 3					
	40	41	27	28	10	11				
c′ g′	2 6	2 10	2 3	2 3	2 2	2 5				
	52	42	43	29	30	12	13			
c′ c″	2 7	2 11	2 4	2 9	2 3	2 6	2 1			
	53	54	44	45	31	32	14	15		
c′ a′♭	2 2	2 6	1 1	2 4	1 2	2 1	1 4	1 5		
	61	55	56	46	47	33	34	16	17	
c′ b′	1 3	2 1	1 6	1 1	1 7	1 4	1 9	1 10	1 5	
	62	63	57	58	48	49	35	36	18	19
c′ a′	2 5	2 9	2 2	2 7	2 1	2 4	1 1	1 2	2 3	2 8
	66	64	65	59	60	50	51	37	38	20	21
c′ b′♭	1 1	2 3	1 4	2 1	1 5	1 2	1 7	1 8	1 3	2 2	1 6

For each block the top number denotes the order of the trial; the first number below denotes whether the first (top heading) or the second (side heading) is the better; the third number denotes the number of demerits assigned.

In each square the *first* figure denotes the order of trials, the *second* the correct preference, and the *third* the amount of demerits (to be explained later), in case of error in the preference.

The sequence of trials is the conventional order in a series of paired comparisons: each two-clang is named by reading the name at the top and the name at the side in two columns that intersect and form a given square. The two-clang given at the top of a column was always sounded first and the one at the side second. The observer was required to express his preference 1 or 2 according as he preferred the first or the second.

The evaluation of the record is based on the ranking of the intervals established in the preceding part, Table II.

The weighting of demerits

In order to arrive at a method of grading the records, an arbitrary scale of demerits was established. The amount of the demerit for each incorrect judgment was computed on the deviation from the norm. One unit of demerit was given for each step of deviation from the norm. Thus, the observer, in his judgment of the pair c′c″-c′g′ signified by recording 2 that the latter interval is the better consonance, he was given a demerit of 1, as c′g′ is in rank removed 1 step from c′c″. If, on the other hand, in his judgment of the pair c′c″-c′e′^{b}, he records his choice as 2, thereby placing the minor third as a better consonance than the octave, the error is more significant, and he is given a demerit of 5. The greatest possible demerit for any one judgment is 11 as in the pair c′c″-c′d′^{b}.

If every answer should be wrong there would be 286 demerits according to Table VI. But this cannot happen because in the long run fifty per cent. of the judgments would be correct by chance, since there are only two possibilities in each case. This reduces the maximum number of probable demerits to 143, which would be the number for one who had *no ability* to appreciate differences in consonance and dissonance and depended entirely upon chance. Such demerit, therefore, is equivalent to 100 ÷ 143, or seven-tenths of one per cent. The records may

then be stated in terms of per cent. of success in conforming to the norm by deducting .7 per cent. from 100 per cent. for each demerit earned. While this weighting is somewhat arbitrary, it does approximate justice to the situation because it is approximately proportional to the magnitude of the error in each case.

Since it would complicate matters to require the observer to follow the order of Table VI in recording, the records were kept in straight columns, the first three containing twenty records each and the fourth six,—one column being left blank between each of these to give space for the marking of error and demerit.

This test was first made on the students in the elementary psychology class in the University of Iowa for the class of '12-'13. After this research had been completed the department furnished similar records for the '13-'14 class; the following year Miss Nesta Williams furnished the records for the '14-'15 class; and Mrs. Esther Allen Gaw contributed the records for the class of '15-'16. Since the publication of the original document has been delayed it is possible to incorporate this large mass of data which have been gathered through the very generous coöperation of the members of the department of Psychology, into one group making 1045 cases.

The norm established.—The general distribution of these cases is shown in Fig. 10, which is based on a number of cases

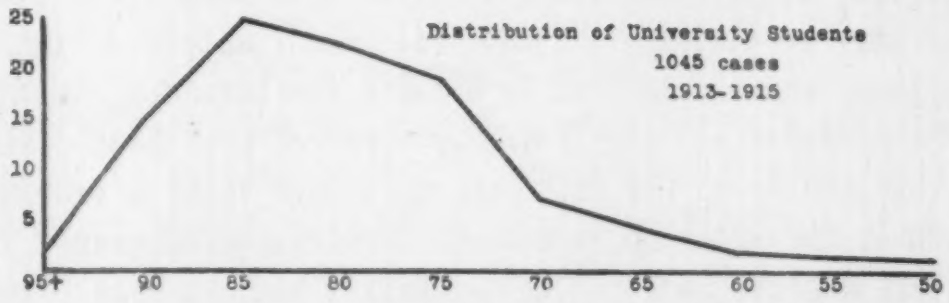

FIG. 10.

sufficiently large taken as a norm. In other words, this curve serves two purposes: it shows how abilities in this act vary and distribute themselves; and it may serve as a norm in terms of which any future record may be interpreted.

Bearing on the order of ranking—The data contained in Part

II bear in themselves a basis for the ranking of intervals which may serve at least as a basis for the criticism of the norm established before. The records for the class of '12-'13 are tabulated to show the total number of preferences for each interval. These are shown in Table VII. In the first column the intervals are given in the order accepted as a norm. (Fig. 9, Table V). The second column contains the per cent. of cases in which each interval was preferred over all with which it was compared. From this the rank of the intervals is indicated in column 3 which shows the deviation of this empirical ranking from the norm.

TABLE VII—*The rank of the intervals in the 1913 class test*

	% of choices	Rank of choices
c′ c″	12.4	2
c′ g′	10.0	5
c′ a′	11.8	3
c′ e′	13.1	1
c′ f′	9.6	6 or 7
c′ a′♭	9.6	6 or 7
c′ e′♭	10.5	4
c′ g′♭	9.8	8
c′ b′♭	6.5	9
c′ d′	3.8	10
c′ b′	2.1	11
c′ d′♭	1.8	12

A glance at this table shows rather satisfactory agreement between the standard order and the empirical order. The orders VIII, IX, X, XI, and XII are the same in both cases, and with these may be counted VI and VII which happen to tie. The deviations occur, therefore, in the first five intervals, chiefly, *i.e.*, in the consonances; and it will be observed that these deviations are due mainly to the fact that the major third is unduly preferred in the empirical rankings. This can be accounted for by the fact that the factor of agreeableness had not been eliminated satisfactorily in the test of this first year. The musical value of the third as compared with the octave must have influenced the students to some extent so that there was a large number of cases in which the major third was preferred to the octave and the fifth. The dropping of the major third in rank will of course raise the octave and the fifth and secure closer agree-

ment with the norm. The same principle is perhaps also operative in tending to throw the middle range somewhat higher than the norm. The intervals $c'e'^{b}$ and $c'g'^{b}$ and $c'b'^{b}$ are rich musical intervals which may have been favored slightly on the same fallacy as that which threw the third off.

It will probably be found that this sort of error was eliminated by more rigid instructions on this point, emphasizing the fact that the preference is an intellectual judgment rather than an appreciation of musical value in the terms of agreeableness.

With this one exception then, the empirical results on the whole tend to confirm the order established in the norm.

The class of '12-13 was required to fill out a questionnaire on musical training in the form given by Professor Seashore in the report of the Committee on "The Standardization of Pitch Discrimination" (25). This questionnaire is designed to secure a measure of general musical ability, musical training, musical environment, and the expression of musical feelings. For the present purposes, the members of the group were ranked on musical training. The Spearman co-efficient R between this musical training and the record for the consonance test is only $R = .02$, which is very remarkable in view of the fact that one should ordinarily expect those who had had musical training to do better than those who had had none. This absence of correlation is of the greatest significance for the value of this test in that it tends to show that it is fairly independent of training and will, therefore, have diagnostic value.

The correlation between perception of consonance and ability to perform, *i.e.*, to sing or play in music is based upon the same questionnaire returns, gave the Spearman coefficient of $R = .06$ which again is an indication of the effect that the perception of consonance is quite independent of training because it would be fair to assume that the degree of correlation would easily be covered up by the principle of selection by which those who have a good ear for consonance would be likely to acquire musical training.

The correlation between pitch discrimination and the perception of consonance by the Pearson product-moments method

was found to be r = .18. This is not so large a correlation as one should really expect and it may be that this is due partly to the fact that pitch discrimination is an immediate sensory experience, whereas the perception of consonance as here tested involves a complex judgment.

The Pearson coefficient for the correlation for consonance and tonal memory is .34 which speaks for the close relationship of these two factors.

Evaluation of the method.—Roughly, the goal of this series of experiments has been reached; we have found a norm for the order of consonances and dissonances and a norm for the distribution of abilities among university students. Both of the fields here opened up must be worked more fully and the two problems are quite distinct. In connection with these experiments, much material has been collected and much is inherent in the records themselves which would throw light on the validity and interpretation of these norms as well as upon the technique of measurement.

But some place must be set for the division of labor; and, since Mrs. Esther Allen Gaw and others have taken up the experimental work in the laboratory from the point at which it was left by the writer in 1914 and propose to carry the analysis of this measurement into finer details, this report must come to a close with the apology that it is only preliminary.

Recommendations Toward a Standard Test

The method of procedure that has been followed, as stated above, can be recommended in giving a test in the perception of consonance, but it is necessary to emphasize certain precautions that must be followed.

In giving the interval, the experimenter must exercise care in the striking of the two keys so as to make the tones equally strong, for the factor of intensity plays a rôle in consonance. The experimenter should practise the combinations and order of pairs until he can play them with ease and without hesitation in an approximately uniform manner.

As has been pointed out by *Seashore* (25), "the tone most

favorable for accurate results" is one of moderate intensity, "the just perfectly, clearly perceptible tone." "It is ordinarily purer than a stronger tone and favors concentration." This precaution in the measurement of pitch discrimination applies as well to the measurement of the perception of consonance. The piano tone is rich and tends to develop many impurities when sounded loudly. These are partially at least if not wholly eliminated through the use of the soft pedal.

Uniformity of duration is important, as the sounding of one tone or note of the two-clang longer than the other proves a distraction and disturbs the observer in his discrimination. It also gives him a clue as to the interval played and changes the nature of the test. The longer duration of one note of the pair tends to change the character of the interval as this note takes a dominant place in the perception, and thus influences the judgment.

The two clangs within the pair should be presented in rapid succession with a constant interval of 1 second, as discrimination of successive stimuli involves the element of memory. A brief interval, therefore, presents the best condition for comparison as "the curve for tonal memory shows that the accuracy of memory falls off very rapidly, immediately after the first second of the interval." (25). A longer interval is necessary between the pairs to eliminate the influence of progression of intervals. Each pair should stand distinctly by itself as a stimulus for comparison and should have no reference to the preceding or succeeding pair. An interval of four seconds accomplishes this end.

The order of presentation of the pairs should be definitely determined by the preparation of a key in which the consonances and dissonances are distributed throughout. This order should also distribute the easy and difficult comparisons, so that the test will be somewhat uniform in difficulty of discrimination throughout sections of the whole test. Such an order prevents undue fatigue at any one point in the test as it conforms in some measure to the double fatigue order, which is a rule of of experimentation in psychological tests. The order also prevents the succession of two-clangs in a definite harmonic series, which must be avoided in the test.

The key given above Table VI was prepared with these precautions in mind, and is well adapted to the measurement of the perception of consonance. If the pairs are to be given in the reverse order, a different key should be made and given separately. The order should not be reversed in the repetitions, as this causes confusion and leads to mistakes in recording.

It has been noted how the element of fatigue may, to some extent, be obviated by the arrangement of the order of presentation. A test of this nature is naturally interesting if conducted in the correct manner, and elicits a continuous conscious effort, which becomes fatiguing if prolonged beyond the endurance of those tested. As has been pointed out by Seashore (25), this fatigue is not a fatigue of the sensory process, but it arises from the concentration of the attention. The attention tends to fluctuate and become distracted when engaged in continuous concentration for any length of time. This is more the case, the greater the conscious effort needed to make an intelligent choice or judgment. The perception of consonance requires this concentration to a very great degree, as it is relatively complex in its nature and requires a keen discrimination and constant application to perceive differences. In this respect, it differs from the mere sensory tests which involve only one factor. As an additional precaution against ennui, or to prevent fatigue, a rest is allowed after each group of ten trials. This divides the one hour recitation time during which the test is given, into three periods, of which the last two, when the element of fatigue is liable to be prevalent, are somewhat shorter than the first period. This tends to equalize the conscious effort, and prevents incorrect judgments due to fatigue.

What has been stated above applies to the sustaining of the attention. In order to secure the attention of the observer, a further precaution is taken in securing the right attitude toward the test, that will stimulate the conscientious application. The experimenter, in giving a class test, must have the energetic and enthusiastic coöperation of all participants. This cannot be overemphasized as an important factor in the measurement of the perception of consonance, which requires such a careful and discriminating application of attention to the details involved.

Summary of Conclusions

Among the conclusions reached in this study the following are prominent:

1. The historical failure to reach an agreement in regard to the rank of consonance and dissonance is due largely to a disagreement as to what constitutes consonance.

2. The perception of consonance is a cognitive process, involving the factors of blending, smoothness, and purity.

3. The order of the ranking of the intervals varies for different qualities of tone. The order has been established for tuning forks, piano, and pipe organ.

4. The uniformity of the distribution curves for different classes tends to show that the factors of the test can be controlled so as to make it well adapted as a class test.

5. A system of weighting errors has been established.

6. The distribution of grades reveals a very great diversity in this capacity among normal observers.

7. The perception of consonance is elemental in a secondary sense in so far as it is based rather on the elemental capacities for pitch discrimination and tonal memory than on acquired musical ability or training.

8. The test in itself is not an adequate measure of musical capacity, but forms one of a series including pitch discrimination, tonal memory, sense of rhythm, tonal imagery, etc. that may be advantageously used for this purpose.

BIBLIOGRAPHY AND REFERENCES

1. Buch, Einar. Über die Verschmelzungen von Empfindungen besonders bei Klangeindrücken. *Phil. Stud.*, 1900, XV, 240.
2. Euler, Leonard. *Tentamen novae theoriae musicae ex certissimis harmoniae principiis dilucide expositae auctore Leonhardo Eulero.* Petroli, 1739 (Chapters II and III).
3. Faist, A. Versuche über Tonverschmelzung. *Zsch. f. Psychol. u. Physiol. d. Sinnessorg.*, 1897, XV, 102-131.
4. Helmholz, H. L. F. *Sensation of tone.*
5. Hohenemser, R. Die Quarte als Zusammenklang. *Zsch. f. Psychol. u. Physiol.*, 1906, XLI, 164-165.
6. Krueger, Felix. Differenztöne and Konsonanz. *Arch. f. d. ges. Psychol.*, 1903, I, 205-275.

7. Krueger, Felix. Das Bewusstsein der Konsonanz: eine Psychologische analyse (Separatdruck) Leipsic, 1905.
8. Krueger, Felix. Consonance and dissonance. *J. of Phil., Psychol. and Scient. Methods,* 1913, X (No. 6), 158.
9. Krueger, Felix. Die Theorie der Konsonanz. *Psychol. Stud.* 1909-1910, V, 294-409.
10. Leibnitz, G. W. *Princips de la nature et de la grace, Leibnitii epistolae ad diversos.* Leipsic; Breitkopf, 240 (Epistel 154).
11. Lipps, Theodore. Tonverwandtschaft und Tonverschmelzung. *Zsch. f. Psychol, u. Physiol.,* 1899, XIX, 1-40.
12. Lipps, Theodore. Das Wesen der musikalischen Harmonie und Disharmonie. *Psychol. Stud.,* 92.
13. Meinong, A. and Witasek, S. Zur Experimentellen Bestimmung der Tonverschmelzungsgrade. *Zsch. f. Psychol. u. Physiol.,* 1897, XV, 189-205.
14. Meyer, Max. Ueber Tonverschmelzung und die Theorie der Konsonanz. *Zsch. f. Psychol. u. Physiol.,* 1898, XVII, 401-421.
15. Meyer, Max. Zur Theorie der Differenztöne. *Zsch. f. Psychol. u. Physiol.,* 1897, XVI, 1-34.
16. Meyer, Max. Unterscheidsempfindlichkeit für Tonhöhen. *Zsch. f. Psychol. u. Physiol.,* 1897, XV; 1898, XVI.
17. Meyers, C. S. and Valentine C. W. A. study of the individual differences in attitudes toward tones. *Brit. J. of Psychol.,* VII, 68-111.
18. Ogden, R. M. A contribution to the theory of tonal consonance. *Psychol. Bull.,* 1909, VI, 297-303.
19. Pear, T. H. Differences between major and minor chords. *Brit. J. of Psychol.,* 1911, IV, 56-94.
20. Peterson, Joseph. Combination tones and other related auditory phenomena. *Psychol. Rev. Monog.,* 1908, IX, 1-136.
21. Preyer, W. Zur Theorie der Konsonanz. *Akustische Untersuchungen,* 1879, 44.
22. Schaefer, K. Psychophysiologie der Klanganalyse. *Ergeb. d. Physiol.,* 1909, VIII, 1-25.
23. Schulze, R. Über Klanganalyse. *Phil. Stud.,* 1898, XIV, 471-489.
24. Schopenhauer, A. The world as will and idea. *Metaphysics of Music,* III, 235.
25. Seashore, C. E. The measurement of pitch discrimination: a preliminary report. *Psychol. Monog.* No. 53, 1910, XIII.
26. Seashore, C. E. The tonoscope. *Univ. of Iowa Stud. in Psychol.,* VI, 1914, 1-12 (*Psychol. Monog.* 69).
27. Stumpf, Carl. *Tonpsychologie.* Leipsic. 1883.
28. Stumpf, Carl. Zum Einfluss der Klangfarbe auf die Analyse von Zusammenklängen. *Beiträge z. Ak. u. Musikwiss,* II, 168-170.
29. Stumpf, Carl. Differenztöne und Konsonanze. *Zsch. f. Psychol. u. Physiol.,* 1905, XXXIX, 269-283.
30. Stumpf, Carl. Konsonanz und Dissonanz. *Beiträge z. Ak. u. Musikwiss,* 1898, I, 91-107.

31. Stumpf, Carl. Beobachtungen über Kombinationstöne. *Beiträge z. Ak u. Musikwiss,* V, 3-133.
32. Stumpf, Carl. Massbestimmungen über die Reinheit konsonanter *Intervalle. Zsch. f. Psychol. u. Physiol.,* 1898, XVIII.
33. Stumpf, Carl. Die Unmusikalischen und Tonverschmelzung. *Zsch. f. Psychol. u. Physiol.,* 1898, XVII, 422.
34. Stumpf, Carl. Über das Erkennen von Intervallen und Accorden bei sehr kurzer Dauer. *Zsch. f. Psychol. u. Physiol.,* 1902, XXVII, 148-186.
35. Stumpf, Carl. Neueres über Tonverschmelzung. *Zsch. f. Psychol. u. Physiol.,* 1897, XV, 280-303.
36. Stumpf, Carl. *Die Anfänge der Musik.* Leipsic. 1911.
37. Sylvester, R. H. Some standardizing tests of Stern tone variators. *Iowa Acad. of Sc. Proceed.,* XVII, 195.
38. Titchener, E. B. Ein Versuch die Methode der paarweisen Vergleichung auf die verschiedenen Gefühlsrichtungen anzuwenden. *Phil. Stud.,* 1902, XX, 382.
39. Valentine, C. W. The method of comparison in experiments with musical intervals and the effect of practice on the appreciation of discords. *Brit. J. of Psychol.,* VII, 118-135.
40. Watt, H. J. Psychological analysis and theory of hearing. *Brit. J. of Psychol.,* VII, 1-43.
41. Wundt, Wm. *Physiologische Psychologie,* II. Leipsic, 1902.

A REVISION OF THE CONSONANCE TEST

BY

ESTHER ALLEN GAW

The consonance test as described by Malmberg in the foregoing article[1] yielded excellent results, but Professor Seashore and Dr. Malmberg both saw the possibility and need of further refinement of procedure. The writer has continued the experimental study of the problem and wishes to report here certain revisions, notably: the elimination of undesirable intervals, the elimination of the system of demerits, and the simplifying of the definition and concept of consonance. A preliminary norm for children has also been established.

After giving the test in the fall of 1915, the records of the three hundred and eight sophomores enrolled in elementary psychology in the University were analyzed for the purpose of determining how the actual order of preference by those tested compared with the norm, what intervals were exceptionally easy or exceptionally difficult and what constant tendencies to error might be discovered.

In Fig. 1, the sixty-six comparisons are so arranged as to show for each comparison first, the number of demerits assigned (heavy solid curve); second, the per cent. of mistakes (broken curve); and, third, the weighted record (dotted curve). The vertical scales for these three curves are respectively, one demerit for each space in the first, ten per cent. of mistakes for each space in the second, and one hundred weighted demerits in the third, the three scales being designated A, B, and C respectively.

The first curve, then, shows simply how many demerits were assigned to a mistake for each pair. The second shows the per cent. of mistakes for each pair. Since there were three hundred and eight cases and each person made one judgment for each pair, the per cent. of mistakes is found by dividing the number of mistakes for a given pair by 308. And the weighted demerits

[1] Malmberg, C. F. The perception of consonance and dissonance (in this Volume).

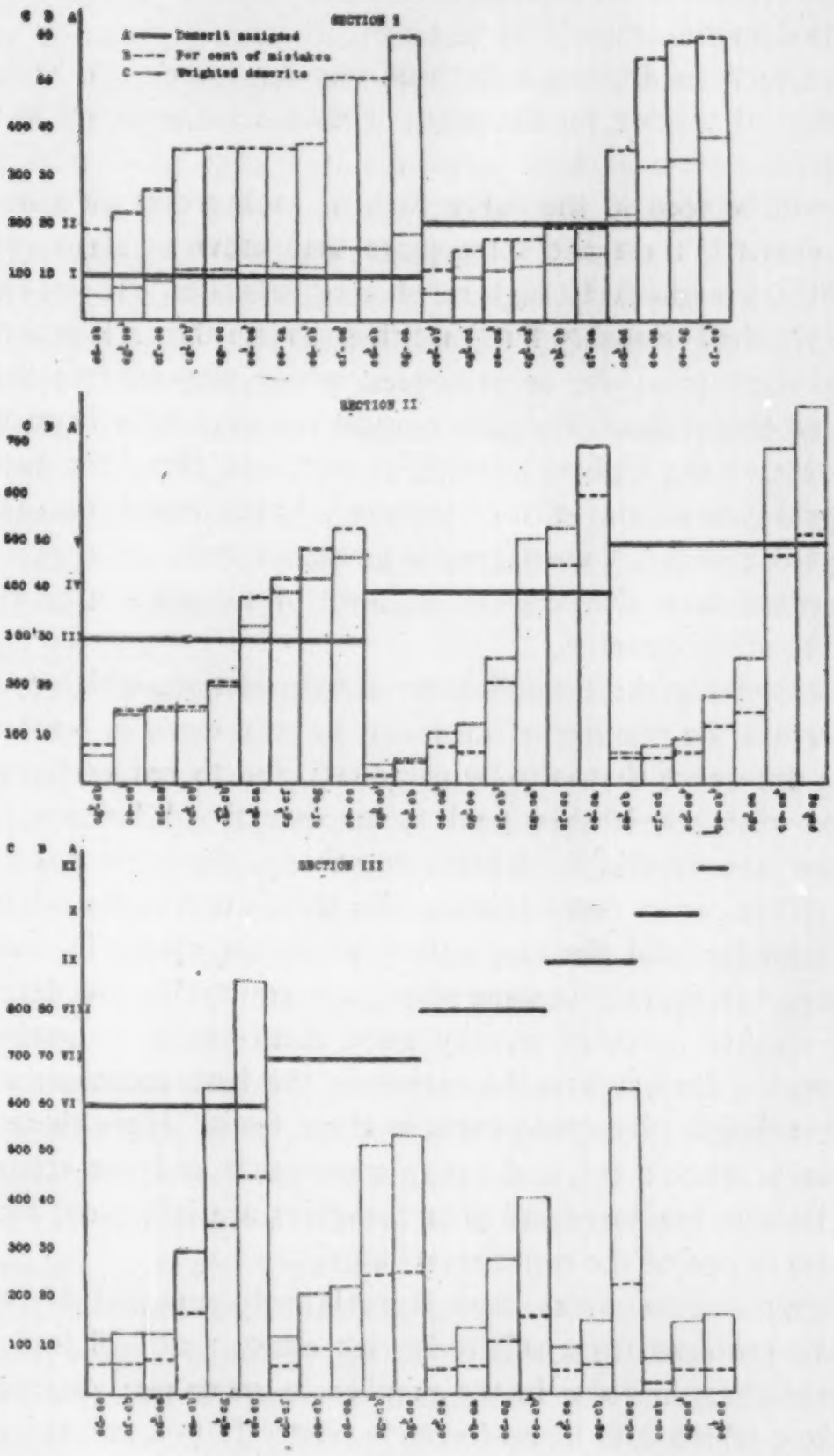
SECTION I
A —— Demerit assigned
B ---- Per cent of mistakes
C —— Weighted demerits
SECTION II
SECTION III

FIG. I

of curve c denote the number of mistakes for each pair multiplied by the demerit assigned to that pair.

The curve is divided into three sections merely for the convenience of placing on the page; it should be regarded as continuous.

It will be seen in the curve that, in each group of pairs for each demerit, there are some pairs that show a relatively low and others a relatively high number of mistakes. If we isolate these we shall probably find that these group themselves according to some principle, or principles, which may explain the respective tendencies. For this purpose we may take from those that receive the highest number of mistakes, three for each of groups 1-4 two from each of groups 5-7; and one from each of groups 8 and 9. A corresponding number from each group of those that show the smallest number of mistakes may represent the other extreme.

In examining these we discover certain motives which operate to increase the number of mistakes, if they work in opposition to the difference that is to be perceived, and to reduce the number of mistakes, if they tend to increase the difference. The former may be said to operate negatively, the latter positively.

A. *Richness of clang quality.* By this is meant the relatively rich, complex, and pleasing quality of certain clangs in contrast to the pure, thin, and, in some respects, harmonically less desirable clang quality of other equally good consonances or semi-consonances. Thus, while the octave is the best consonance, the major third is given that rank in these tests. Here there is a conflict between a relatively pure consonance and one relatively rich, because mediocre and poor observers are influenced by this richness in one of the members of a pair.

Throughout our series, then, if a relatively pure and thin clang is to be preferred to a relatively rich clang, that will lead to a corresponding increase in the number of mistakes; whereas, if the clang which is to be preferred is relatively rich and the other relatively pure, that will lead to a corresponding decrease in the number of mistakes. In the former case, this motive operates negatively; in the latter, positively.

B. *Similarity of dissonance or semi-dissonance quality.* It is easier to compare two dissonances of the same kind such as c′d′-c′d′♭ than to compare two dissonances of different kinds such as c′d′-cb♭, regardless of their order of rank, because it is easier to compare two things of the same kind than two things of different kinds. Like motive A, this may operate either negatively or positively. If the comparison is between dissonances of different kinds, it will increase the number of mistakes, whereas if it is a comparison of dissonances of the same kind, it will tend to lessen the mistakes.

While, from one point of view, this is a motive for constant error, it is, from another point of view, a condition for an actual difference, either positive or negative. That is, it makes certain comparisons actually easier or more difficult (according as it acts negatively or positively) than the difference in rank in the norm would indicate.

C. *Conspicuous roughness.* Roughness in the dissonances is by far the easiest mark of dissonance. c′d′, *e.g.,* is always conspicuous for its roughness; and, wherever that interval is involved, it is relatively easier to detect it than to detect a dissonance of the c′b′ type. This again may operate either positively or negatively. If the poorer clang is c′d′, this motive acts positively and makes the judgment easier; whereas, if c′d′ is the better clang, it acts negatively and makes the decision more difficult.

Like motive B this is also a condition of actual difference in the ease or the difficulty of recognizing a given interval in which it operates. In that respect motives B and C differ from motive A, which is merely a motive for constant error; they are also conditions of actual difference.

The cases grouped under these motives are really merely illustrative, because not only are there more instances in which two motives coöperate, but it is a practical certainty that there are numerous cases in which two or more motives conflict and result in an interference which reduces or eliminates the net error. There are also many cases in which more than two motives for constant error are present. Those mentioned are, however,

probably the most important, because they go far to account for the deviation of the empirical order from the true order of rank.

There are undoubtedly other motives present, but for a rough preliminary analysis, we may arrange the intervals which were excessively high or excessively low in mistakes, as indicated in Fig. 1, under the negative or positive aspect respectively of each of these three motives as follows:

After each interval is indicated the per cent. of mistakes of each pair.

A. *Richness of tone quality*

Negative
Excessive Mistakes

c′ g′-c′ a′	(60)
c′ f′-c′ e′^{b}	(60)
c′ e′-c′ g′	(56)
c′ c″-c′ a′	(53)
c′ e′-c′ c″	(50)
c′ g′-c′ f′	(47)
c′ e′^{b}-c′ a′	(57)
c′ f′-c′ c″	(43)
c′ g′-c′ a′^{b}	(40)
c′ g′-c′ e′^{b}	(50)
c′ c″-c′ a^{b}	(45)
c′ e′^{b}-c′ c″	(46)
c′ g′^{b}-c′ c″	(25)
c′ g′-c′ b′^{b}	(24)
c′ c″-c′ d′	(22)
c′ f′-c′ a′^{b}	(49)
c′ f′-c′ g′	(45)
c′ c″-c′ b′^{b}	(16)

Positive
Minimum number mistakes

c′ g′^{b}-c′ e′^{b}	(22)
c′ e′^{b}-c′ b′^{b}	(11)
c′ g′^{b}-c′ b′	(7)
c′ a′^{b}-c′ b′	(4)
c′ e′^{b}-c′ b′	(5)
c′ f′-c′ b′	(6)
c′ e′-c′ b′	(9)

B. *Similarity in tone quality*

Negative
Excessive mistakes

c′ d′^{b}-c′ b′	(51)
c′ g′^{b}-c′ g′	(34)

Positive
Minimum number mistakes

c′ e′-c′ a′^{b}	(17)
c′ e′-c′ e′^{b}	(15)

C. *Roughness*

Positive
Minimum number mistakes

c′ d′-c′ b′	(19)

A-C *Musical Quality and Roughness*

Positive
Minimum number mistakes

c′ g′^{b}-c′ d′	(15)
c′ g′^{b}-c′ d′^{b}	(4)
c′ e′-c′ d′	(6)
c′ d′^{b}-c′ e′	(4)
c′ f′-c′ d′^{b}	(6)
c′ a′-c′ d′^{b}	(4)

If we examine the curves showing the demerit assigned (heavy solid curve, Fig. 1) and the result of weighting the demerits for each pair (dotted curve, Fig. 1) we find that the actual order of difficulty is not inversely proportional to the distance in the order of rank, as is assumed in the system of demerits. This is due primarily to the constant errors which have been noted, of which the chief one is the tendency to be influenced by motive A, musical agreeableness. But, as was pointed out above, motives B and C are also causes or conditions of *actual clearness of difference*. The norm is based on the actual order of consonance-dissonance. These notions will, therefore, operate to make the empirical order of rank legitimately different from that of the norm, because it makes certain pair-differences either more or less clearly perceptible.

Malmberg found that the empirical rank in 1913 did not correspond to the norm. The same was found for these records of 1915 which were taken by the same method. A comparison of the two deviations from the norm is shown in Table I, which gives, first, the order of rank in the norm as established by Malmberg; second, the empirical order as determined by the per cent. of choices in the 1913 records; and, third, the empirical order of rank as established on the basis of the per cent. of choices in 1915 records.

TABLE I

Malmberg norm	*Empirical order of 1912-13*		*Empirical order of 1915*	
	Rank	% of Choices	Rank	% of Choices
1 c′ c″	2	12.4	3	11.4
2 c′ g′	5	10.0	7	10.1
3 c′ a′	3	11.8	2	12.2
4 c′ e′	1	13.1	1	12.7
5 c′ f′	6 or 7	9.6	4	10.8
6 c′ a′^{b}	6 or 7	9.6	6	10.5
7 c′e′^{b}	4	10.5	5	10.7
8 c′ g′^{b}	8	8.8	8	6.7
9 c′ b′^{b}	9	6.5	9	5.8
10 c′ d′	10	3.8	10	4.1
11 c′ b′	11	2.1	12	2.0
12 c′ d′^{b}	12	1.8	11	2.1

In this table it will be observed that the principal deviation is among the consonances, but extends somewhat downward

through the semi-consonances,—whereas the dissonances and one semi-consonance maintain their rank according to the norm, except for a very slight reversal of the last two in 1915, which is negligible. In these deviations from the norm we notice the same general principle that was designated above as motive A, to the effect that there is a tendency to give undue preference to the musically rich and harmonically desirable intervals and that, therefore, the relatively pure and thin consonances are forced down in rank. Thus, the octave is forced down to rank two in one group and rank three in the other. The perfect fifth is forced clear down to the fifth rank in one group and to the seventh rank in the other; but the sixth maintains its rank in one group and is raised one point in the other.

The major third shows the most notable change in that it is given first rank in both groups. The fourth practically maintains its balance, falling slightly in one group and rising slightly in the other. The minor sixth maintains its rank, and the minor third is raised three points in rank in one group and two in the other. That is, the most notable advances are the major third and the minor third which are both conspicuous for the very rich tone quality of the two-clang, and the most notable drop is that of the fifth and octave which, though pure, lack that very richness which gives preference to the ranks immediately below.

The greatest source of error in the test is, then, the failure to adhere rigidly to the definition of consonance by allowing musical agreeableness to play a rôle in the decisions. This factor must, therefore, be further controlled in future tests.

The analysis thus far then has shown that, if the system of weighting is to be continued, it should be readjusted empirically on the basis of records in which the above motives for constant error have been eliminated in so far as possible. To this problem we shall return later.

The analysis has also shown that certain intervals are undesirable either because too easy or too difficult, as the result of the operation of causes for constant error. The natural remedy for that is to eliminate these, and limit the material to

those pairs for which these sources of error are relatively absent.

In order to reduce the number of factors to the minimum and to give the widest range of choice from the material available, it was decided to select only eleven pairs—one for each degree of difficulty as represented by the original system of demerits, which is based on the order of rank in the consonance-dissonance norm. There is manifestly a great advantage in using a small number of pairs, other things being equal, because these may then be examined in fuller detail and be better controlled, aside from the fact that they are the best from the larger group of the original test.

The following points were kept in mind (a) Each interval in the octave shall appear at least once. (b) Since Malmberg's order differs from the customary order in harmony in the case of the intervals, c′a′ and c′f′, no pairs shall be chosen which would be affected by that reversal. (c) Since it is very difficult to judge an interval on the criterion of purity alone, and, since the tendency is to judge a rich clang better than a pure clang, because it is more pleasing, all comparisons which rest primarily upon the criterion of purity, or richness, according to Malmberg's classification of the criteria (Table VI) shall be avoided. For example, we do not use the comparison c′c″-c′e′ nor c′g′-c′a′. (d) In addition to the above four restrictions, the extremely difficult and the extremely easy comparisons, according to the showing in Fig. 1, shall be eliminated where consistent with the observance of the other restrictions. (e) These shall be comparisons of consonances with consonances, semi-consonances with semi-consonances, and dissonances with dissonances, and each of these three with each of the other two.

Reference to the Malmberg Table VI, showing the pairs arranged according to demerit, will help in keeping the following choice of intervals clear.

The following is a list of the pairs finally selected, with statement of the reasons for the choice:

Octave and minor second (c′c″-c′d′^{b})—eleven demerits.

This was the only possible interval with that number of demerits.

Octave and major seventh ($c'c''$-$c'b'^{\flat}$)—*ten demerits.*

The only other interval with this number of demerits is a comparison of a smooth and pure consonance with a very rough dissonance.

Perfect fifth and major seventh ($c'g'$-$c'b'$)—*nine demerits.*

One of the other possibilities was c′d′-c′c″ which was discarded because of the abnormal number of mistakes. The other possibility, c′ a′-c′ d′$^{\flat}$, was discarded because it was a comparison of a pure consonance and a rough dissonance.

Major sixth and major seventh ($c'a'$-$c'b'$)—*eight demerits.*

This seemed the best of the four possibilities, because it is a comparison of a clear consonance with a clear dissonance.

Major sixth and major second ($c'a'$-$c'd'$)—*seven demerits.*

Of the four other possibilities, one contains a very rough dissonance, one contains c′ b′ which we have already used three times, and the other two mixed comparisons are found to be hard to judge.

Perfect fifth and diminished fifth ($c'g'$-$c'g'^{\flat}$)—*six demerits.*

Four of the other pairs contain c′ e′, c′ e′$^{\flat}$, c′ a′, and c′ a′$^{\flat}$, respectively, which are so often incorrectly judged on account of their rich and pleasing qualities. The other one possibility contains c′ b′ which cannot be used again. The result is that we have an interval which is relatively more difficult that its rank would indicate.

Major third and minor seventh ($c'e'$-$c'b'^{\flat}$)—*five demerits.*

There were six other possibilities. Two of these were c′ e′$^{\flat}$-c′ g′, and c′ a′$^{\flat}$-c′ c″, each of which is a comparison of a very pure consonance with a rich consonance which has been found too hard to judge considering its high demerit. Two others, c′ f′-c′ d′, and c′ g′$^{\flat}$-c′ a′, are too much of the type of the six pairs already chosen. The other possibility contains c′b′, which has already been used three times.

Perfect fifth and minor sixth ($c'g'$-$c'a'^{\flat}$)—*four demerits.*

This is a type of comparison that ought to be included, that of a pure consonance with a rich consonance, and, since the demerit is not too high, it seems best to use this pair here.

Major third and minor third ($c'e'$-$c'e'^{\flat}$)—*three demerits.*

This is a comparison of two rich consonances, and therefore relatively easy, but c′ a′-c′ a′$^{\flat}$, of the same order, would have been a better choice.

Major seventh and minor seventh ($c'b'$-$c'b'^{\flat}$)—*two demerits.*

This is a comparison of two dissonances, a type which must be included.

Perfect fourth and major third ($c'f'$-$c'e'$)—*one demerit.*

This was chosen in order to include the perfect fourth, but it is really too easy.

To determine in a preliminary way the effect of the use of these eleven pairs alone, the record on these intervals was taken out from the complete record of the sixty-six pairs in the 1915

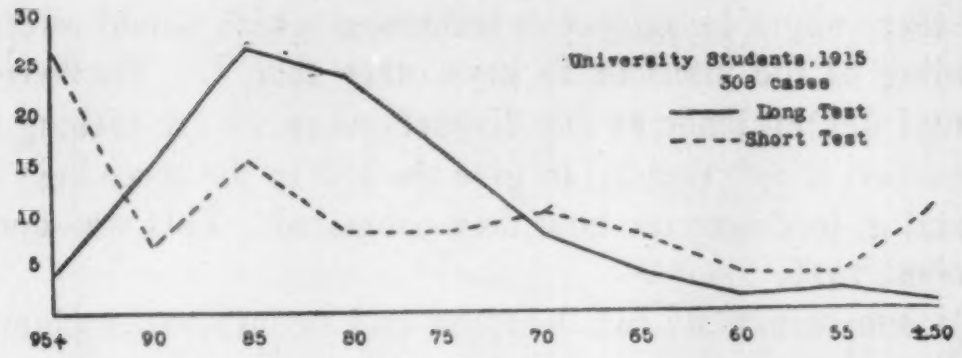

FIG. 2

test. Fig. 2 shows a comparison of the distribution according to the 66 unit test (the original) and the 11 unit test, by the method of weighted demerits. The curve shows that the 11 unit test is very much easier than 66 unit test. This difference in difficulty is, of course, due (1) to the elimination of the extremely difficult pairs, and (2) to the elimination of the principal motives for constant error.

A serious question, therefore, arises as to whether the test in this new form is too easy. It could, of course, be readily changed to any desired degree of difficulty without reintroducing seriously any of the sources of error which have been eliminated, but, in view of the desirability of having the norm for the adults comparable with the norm for children whose records run lower, it was decided to keep the material with the degree of difficulty here found. However, the new method of evaluation introduced below makes the test even with this material more difficult again.

So far the intervals have always been considered as occurring in the key of C. There is a decided advantage in varying the key to avoid the monotony of using the same notes, c′, c″, throughout. And there is no apparent reason why it should not be given in other keys, except perhaps that the tempered scale used on the piano would cause the intervals to have a different consonance value. According to a law developed by Stumpf[2] there should be no difference in the consonance of the various intervals in different keys even in tempered scale.

[2] Stumpf, Carl. Beiträge zur Akustik und Musikwissenschaft, III, Table IX. Leipsic 1898.

But there might be subjective influences which would affect the standing of the intervals in keys other than C. Furthermore, Stumpf did not make as fine distinctions as we are making here. Therefore, it was decided to give the test in the three keys of c, b^b, and g, to determine how they compared. This was done in October, 1916.

To one section of two hundred and twenty-two sophomores the list of eleven pairs was given four times in the key of c, and four times in the key of b^b, and four times in the key of g.

The result is shown graphically in Fig. 3 in terms of the

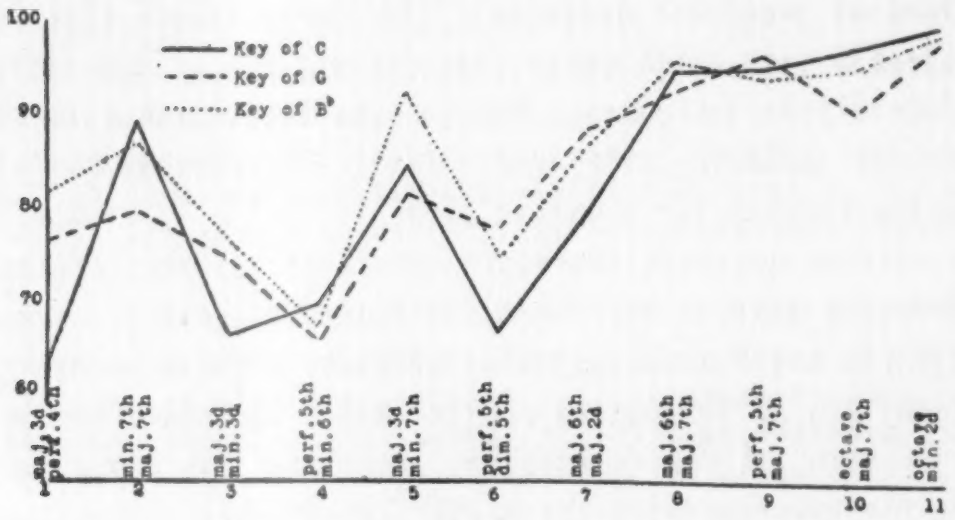

FIG. 3

percentage of correct judgments for each pair in each key. As will be seen, the correct judgments are approximately the same for each of the three keys used, which corroborates under any conditions, and to a serviceable degree, the principle laid down by Stumpf.

Fig. 3 also shows certain fluctuations in the degree of difficulty for different pairs, which are strengthened by the fact that they tend to agree for the three different keys. These may be interpreted in the light of the previous discussion.

The selection of the eleven pairs from the 1915 test yielded, of course, only one judgment for each interval, which is too little to be reliable. In a regular test, the pairs must be repeated as many times as possible to get a large number of trials.

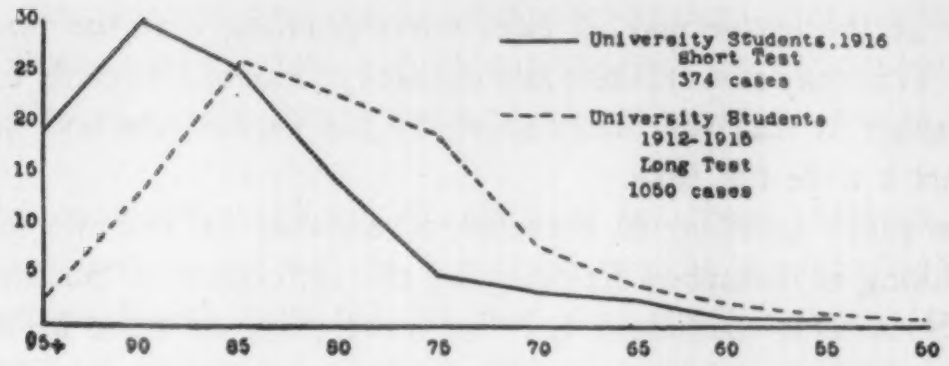

FIG. 4

Fig. 4 shows the distribution of this 1916 group of records as compared with the distribution of Malmberg's 1050 cases in the original form of the test. It will be seen that the difference between these two distributions is analogous to that found in Fig. 2. Fig. 4 is undoubtedly more reliable, because the dotted curve is based upon a larger number of cases, and the solid curve is based upon eighty-eight judgments as compared with eleven in Fig. 2.

If we should regard a selected group of pairs as, *e.g.*, the eleven intervals here selected, each played in three or four keys, as a single unit, it would simplify procedure very much to do

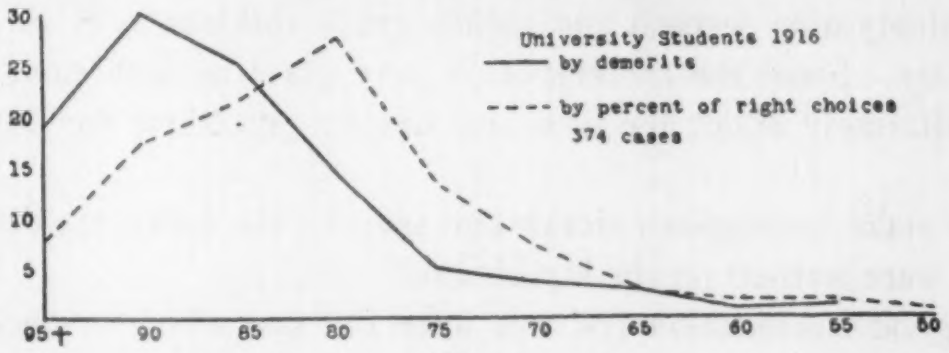

FIG. 5

away with the system of ranking errors by demerits, and use simply the per cent. of right judgments as a measure of ability. This is, of course, the reciprocal of the number of mistakes.

In Fig. 5, this method is compared with the original method of grading by weighted errors. The disparity between the two curves shows that the method of demerits tends to yield a higher ranking than the method of mistakes, as the latter curve falls

below at the better end, it rises correspondingly at the poorer end. This may not indicate any disparity; indeed, it points to an advantage in that the distribution by the former method made the test a trifle too easy.

The main question to consider is whether or not the order of ranking is disturbed seriously by the difference in tabulation. The Spearman coefficient for the correlation of rank between

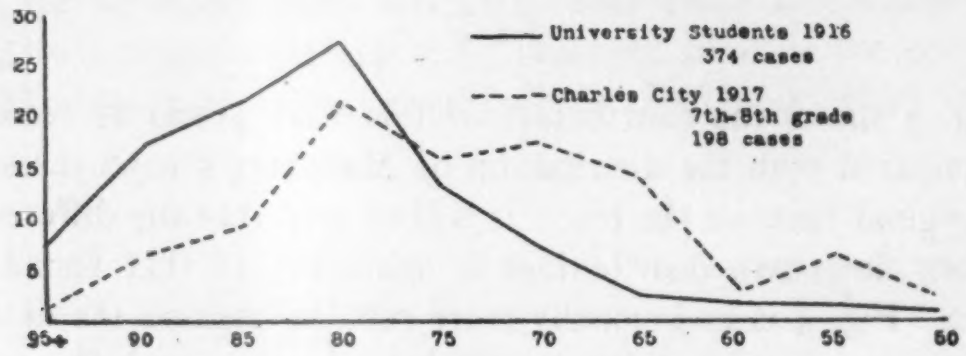

FIG. 6

the two methods for the three hundred and seventy-four cases of University students is R = .99, p.e., .02. This correlation is higher than the actual facts would indicate. But the Spearman coefficient of rank between the two methods for one hundred and ninety-nine seventh and eighth grade children is R = .89. p.e., .03. Since the correlation is very good in both cases we unhesitatingly adopt the new and simpler procedure for future use.

To make instructions simple and specific, the following definitions were written on the blackboard:

A good combination (of two tones) is one which is *smooth,* in which the two tones *blend,* and seem to *agree* and *fuse together* into one.

A bad combination (of two tones) is one which is *rough,* in which the two tones do not *blend,* and do not *seem* to *agree nor fuse together* into one.

Thus far all our efforts to establish norms have been limited to records of adults. All records taken on children have been evaluated temporarily in terms of the adult norms. We are at present engaged in establishing norms for children. Merely as a preliminary suggestion of what changes may be expected for chil-

dren we may use here the records of a group of one hundred and ninety-nine children, which comprise all the children in the seventh and eighth grades in the public schools of Charles City, Iowa.

The comparison of the distribution for these children with the distribution for adults is shown in Fig. 6, in which the curves are based upon the per cent. of correct judgments irrespective of demerits. Here is then a clear-cut picture of the ability of children of this age as compared with adults.

The following table is a comparison of the percentile rank of adults and children according to the per cent. of correct judgments. The first and second columns give the per cent. of correct judgments which an adult and a child respectively would have to make in order to get the percentile rank set opposite in the third column. It will be seen that in order to be average in the perception of consonance, or to get a percentile rank of 50, an adult needs to make 70% of correct judgments, while a seventh or eighth grade child must make 54% of correct judgments.

Adult	*Child*	*% Rank*
100	97	100
86	79	90
82	69	80
77	66	70
75	60	60
70	54	50
66	48	40
61	43	30
54	33	20
46	21	10
9	0	0

The establishment of norms for children and their evaluation for different ages represents a large task in the midst of which we are now engaged. In view of the strenuous nature of the test, one must be impressed with the relatively small difference between the achievement of adults and children. This fact becomes even more impressive upon examination of individual evidences. Thus, many of the children make as good a record as the best of the adults, and the disparity between the showing of children and adults comes largely from the incompetency of the comparatively dull children.

THE COMPARATIVE SENSITIVENESS OF BLIND AND SEEING PERSONS

BY

CARL E. SEASHORE AND T. L. LING[1]

The blind have unquestioned superiority in their ability to use their touch, hearing, and other senses for guidance. From observation of this, the conclusion is generally drawn that the sensitivity of their sense organs is correspondingly heightened.

"Through years of experience in the laboratory, the conviction has gradually grown upon me that a more radical distinction should be made between sensitiveness and ability to use a sense; *i.e.*, between inborn sensory capacity and acquired ability or skill. From time to time I have taken the opportunity of comparing my own sensitiveness in touch and hearing with that of blind persons distinguished for ability in guiding themselves by hearing and touch; and in no case did I find that the blind persons possessed any significant superiority to myself in sensitiveness to touch and hearing, although some of the blind persons experimented upon were noted for their wonderful performances through hearing and touch."[2]

We finally found an opportunity to put the matter to an accurate and crucial test. Sixteen blind students from the Iowa College for the Blind at Vinton, Iowa, and fifteen pupils from the Iowa City High School were selected as cases for comparison. The ages of the blind ranged from sixteen to twenty-six and of the seeing from fourteen to nineteen. In selecting the blind, only those were chosen who had been totally blind for more

[1] The writers acknowledge their great obligation to Superintendent Eaton for courtesies and coöperation at the School for the Blind and to Dr. Mabel C. Williams for taking responsible part in the making of the measurements, and taking the case histories of the blind children.

[2] Seashore, Carl E., Elemental Tests in Psychology, *J. of Educ. Psychol.*, 1916, 7, 81.

than five years, who were generally otherwise both mentally and physically sound, and were now in high school grades. This resulted in the selection of the ablest pupils in the school. The seeing pupils were taken at random, being selected by the principal on the ground of convenience of schedule. On each of these the following six measurements were made.

List of tests

1. *Localization of sound.*—Discrimination for the direction of sound was measured in terms of the angular displacement in the horizontal plane directly in front of the observer. A click in a telephone receiver served as stimulus. The receiver was mounted at one end of a two-meter bar which was pivoted at the middle and carried a pointer running over a scale of degrees at the other end. This apparatus was mounted at such distance in front of the observer that the receiver moved at the distance of one meter from the center of the head. The click was produced by the rapid make and break of the circuit through the mercury key. This whole arrangement proved very accurate and convenient, inasmuch as the angular displacement at this point and in this direction is very small. After a preliminary determination of the threshold one hundred trials were made by the method of constant stimuli.

2. *Discrimination for the intensity of sound.*—This was measured by determining whether the second of two consecutive tones which differed in intensity only, was stronger or weaker than the first. The tones were produced by a one hundred vibration electro-magnetic fork in circuit with the audiometer[3]. A standard current of 1.2 volts was used and No. 30 on the audiometer was taken as the basal unit, the compared tones being either stronger or weaker. This provides a tone clearly audible and yet not unpleasantly strong. The records were kept in terms of the units of the audiometer, which represent approximately equal psycho-physic steps. After preliminary determination of the threshold, one hundred trials were given by the method of constant stimuli.

[3] Seashore, New Psychological Apparatus, *Iowa Studies,* Vol. II, pp. 153-163.

3. *Discrimination for lifted weights.*—A set of bottles forty millimeters in diameter and ninety-five millimeters high was arranged as follows: 3 standards at 80 grams each and 12 compared weights yielding a series—81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 92, and 95 grams. The observers were blindfolded and the standard weights were changed in order to eliminate the temperature error. The weights were arranged with ample space on a small revolving table so that any desired weight could be swung into position without noise. After the preliminary determination of the threshold, one hundred trials were made by the method of constant stimuli.

4. *Discrimination for passive pressure.*—A modified form of the Stratton pressure balance[4] was used. This was built of aluminum which was lighter and therefore more sensitive than the original. By this form of the apparatus it was possible to increase or decrease the standard pressure instantaneously without the slightest jar. The standard pressure was five grams upon the outer side of the third phalanx of the first finger. The pressure location of this was changed slightly for successive trials and the least perceptible increase or decrease in weight was determined, one hundred trials being made by the method of constant stimuli after the preliminary threshold had been determined.

5. *Discrimination for active pressure.*—This was measured by determining under how many thicknesses of paper the position of a wire hair could be located by active exploration with the tip of the first finger. The apparatus consisted of a plate glass disc 10 cm. in diameter, mounted so as to revolve freely. In place of the hair a copper wire, No. 38 B & S, which is about the diameter of the average human hair, was stretched across the plate glass and with ends firmly fastened. Discs of paper cut from Old Veda Bond 17 x 22, twenty pounds to the ream, were placed over this wire, and the blindfolded observer was required to locate the direction of the wire through touch. Starting with ten sheets, five sheets were added for each trial until

[4] Stratton, Über die Wahrnehmung von Druckänderungen bei verschiedenen Geschwindigkeiten, *Philos. Stud.*, Vol. 12, 1896, pp. 525-586.

a failure occurred; then a single sheet was taken off or added until the greatest thickness through which the wire could be located had been determined. The record is therefore given in terms of the number of sheets of paper covering the wire. To one who has not performed this test, the actual results will seem to be very wonderful achievements, as the records show.

In this test much depends upon economic procedure in the hunting. The observers were therefore carefully taught how to proceed: *e.g.,* to run tip of finger from left to right and, if the wire was not found in that stroke, to run it in a line at right angles to the first; the wire being found by such economic procedure, it was comparatively easy to trace it. If it was not found in the first two strokes, these cardinal movements were continued until the observer acknowledged failure.

6. *Tactual space.*—This consists in the determination of the two-piont threshold in the transverse direction on the tip of the first finger and in the transverse direction on the inside of the forearm, 5 cm. above the wrist.

It is unfortunate that no more than one hundred trials could be made in each of these tests, where the method of constant stimuli was employed, but it must be remembered that much time, more than was required for the actual measurement, was spent in making preliminary determination and securing adaptation which is necessary for tests of this sort. Indeed the elemental character of the tests often hinges largely on full knowledge of all factors involved and skill in preparing the observer so as to eliminate sources of subjective and objective variables.

Case histories: 1. The blind

Case 1. Age, 16, female. Blindness began three days after birth. Can only distinguish light from darkness. Attractive, bright, quick, responsive, and breathlessly attentive, but easily fatigued; reads point with either hand at conversational rate; likes typewriting, plays the piano well, and sings; can hem by hand but does not use a sewing machine; taste and smell are good; has no difficulty in finding her own way on the premises.

Case 2. Age, 26, female, single. Cause of blindness *iridocyclitis;* trouble came on gradually from ten to nineteen. Can tell where windows are in a room; healthy and cheerful, but slow; responsive and tried hard in the tests;

reads point slowly with right hand and follows with left; studied typewriting for a while but dropped it; can sew by hand or machine, crochet, knit, tie, and do housework; has learned basketry, weaving, and broom making; has studied piano two years; taste and smell good.

Case 3. Age, 25, male, single. *Congenital cataract.* Can tell light from darkness; taste and smell good; relies much on hearing; is of a nervous temperament, unattractive, but quiet; critical and sure of himself; reads point slowly with left index finger; uses typewriter, studied piano nine years; profession, piano tuning; also makes brooms, ties, and weaves.

Case 4. Age, 19, female. Blindness due to *leucoma adherens* from *blennorrhea neonatorum;* commenced shortly after birth; can locate windows with left eye; can sort laundry by smell; makes much use of hearing; was fidgety and sighed frequently during the test; bright and alert; reads point rapidly with both hands; plays piano well; sews either by hand or machine; skillful in basketry and fancy work.

Case 5. Age, 23, female, single. Blindness due to *sympathetic ophthalmia;* ran a knife into her right eye when a child and failed to get proper treatment; can tell daylight from darkness with her left eye; goes about alone in usual places but never on the street; reads point with both hands at conversational rate; sews, ties, crochets, etc.; tries to do all kinds of house work; plays the piano well; fine young woman, neat and pleasing, quiet manner.

Case 6. Age, 18, male. Blindness due to *phthisis bulbi* from *blennorrhea neonatorum;* pleasing appearance, alert, attractive, but fatigues readily; reads point with left index finger at conversational rate and anticipates with right; uses typewriter well; is musical and sings well; guides himself in the environment without difficulty.

Case 7. Age, 19, male. Blindness due to *phthisis bulbi.* Stuck a shoe awl into right eye when nine years old and rapidly lost sight of both eyes a year later. Guides himself in environment with skill and ease; taste and smell good; basket weaving, broom making; carpet weaving; reads point slowly with left index finger; plays piano and violin a little; likes typewriting.

Case 8. Age, 19, male. Blindness due to chronic *conjunctivitis;* began when six months old. Can locate windows and see the light of the sun but not the sphere itself; taste and smell good; cheerful, reliant, and mentally probably a little above the average; reads point; tunes piano; takes typewriter apart just for the fun of it; has a work shop and whole set of tools at home; built a house seven feet square twenty feet up in a tree two years ago.

Case 9. Age, 19, female. *Atrophy of optic nerves* due to accident at age of fourteen. Intelligent and responsive and tried hard in the tests; uses typewriter; is skillful in basketry and fancy work.

Case 10. Age, 17, male. Blindness due to chronic *iridocyclitis* caused by *spinal meningitis* at six months. Lost sight suddenly at age of twelve; taste and smell good; reads point with either hand at conversational rate; can work in the garden, but not hoe; makes baskets and hammocks.

Case 11. Age 19, female. Blindness due to *sympathetic ophthalmia;* ran scissors into left eye when four years old; right eye failed gradually. Can see light, locate windows, and avoid obstacles; ordinary intelligence; reads

point with left index finger at conversational rate but not with right hand at all; plays the piano; uses typewriter, sews, and does fancy work.

Case 12. Age, 19, male. Blindness due to *maculae corneae from blennorrhea neonatorum.* Goes about the city alone; learns new towns quickly; conscious of using hearing extensively; responsive and attentive; boastful; reads point well; plays pipe organ, piano, and violin; gives concerts.

Case 13. Age, 20, female. Blindness due to *congenital atrophy* of optic nerves. Taste and smell good, but does not use them much; good orientation; nervous; reads point with right index finger, left following; uses typewriter, sews, and does fancy work.

Case 14. Age, 21, female. Blindness due to *phthisis bulbi* from *blennorrhea neonatorum.* Can detect light with the left eye; reads point and braille with either hand but faster with the right; plays piano and sings well; sews and does fancy work, and house work.

Case 15. Age, 25, female, single. Blindness due to *phthisis bulbi* from *blennorrhea neonatorum.* Cannot detect light; taste and smell good; responsive, good application; reads braille better than point; sews and does fancy work and light house work.

Case 16. Age, 20, female. Blindness due to *phthisis bulbi* from *blennorrhea neonatorum.* Can detect light and sometimes colors and outlines of persons; bright, very attentive; reads point with right index finger rapidly and is considered one of the best; uses typewriter, plays piano, and pipe organ, sews, and does house work.

2. The seeing

Case 1. Age, 18, male, junior. A good student.
Case 2. Age, 14, male, freshman. A medium student.
Case 3. Age, 17, female, junior. A medium student.
Case 4. Age, 16, male, freshman. A medium student.
Case 5. Age, 21, male, senior. A superior student.
Case 6. Age, 17, male, freshman. A medium student.
Case 7. Age, 17, male, freshman. A medium student.
Case 8. Age, 19, female, junior. A good student.
Case 9. Age, 18, female, senior. A very bright student.
Case 10. Age, 17, male, senior. A medium student.
Case 11. Age, 18, male, senior. A medium student.
Case 12. Age, 18, female, senior. A medium student.
Case 13. Age, 15, male, sophomore. A medium student.
Case 14. Age, 15, male, freshman. A very poor student.
Case 15. Age, 15, male, sophomore. A good student.

TABLE I—*The blind*

Cases	Localization of Sound	Intensity of Sound	Lifted Weight	Pressure	Touch	Tactual Finger	Space Arm
	Degrees	*Units*	*Grams*	*Grams*	*Sheets*	*m. m.*	*m. m.*
1	1.4	1.5	3.1	.64	25	1.5	20
2	3.0	2.0	5.6	.32	32	1.8	
3	5.2	.5	4.0	.63	31	1.3	51
4	1.4	1.3	3.2	.18	21	1.6	30
5	1.3	1.4	4.8	.66	35	1.5	35
6	1.8	.8	4.3	.46	30	1.1	21
7	2.1	2.0		.34	30	.6	27
8	2.0	.8	2.5	1.31	35	1.0	12
9	4.8	1.4	5.2	.66	29	1.1	47
10	2.0	.7	3.0	.38	10	.9	20
11	2.1	1.8	6.6	.38	25	.8	5
12	1.3	.7	1.7	.35	38	1.3	24
13	1.4	1.1	4.0	.22	19	2.3	20
14	1.2	1.2		.51	18	1.3	16
15	1.6	.9	4.0	.70	15	.8	30
16	1.6	.8	4.0	.25	20	2.0	24
Median	*1.6*	*.9*	*4.0*	*.38*	*25*	*1.1*	*24*
Ave..	*2.1*	*1.2*	*4.3*	*.50*	*26*	*1.3*	*25*
M.V.	*.8*	*.4*	*.9*	*.20*	*6.7*	*.4*	*8.6*

TABLE II—*The seeing*

Cases	Localization of Sound	Intensity of Sound	Lifted Weight	Pressure	Touch	Tactual Finger	Space Arm
	Degrees	*Units*	*Grams*	*Grams*	*Sheets*	*m. m.*	*m. m.*
1	1.5	1.0	2.9	.13	41	.6	26
2	4.4	1.6	5.2	.98	14	2.1	19
3	2.0	1.0	2.4	.87	18	1.0	35
4	2.2	1.0	4.6	.50	28	1.7	38
5	2.2	.7	3.0	.35	40	1.6	33
6	1.4	1.6	3.3	.72	47	1.2	26
7	2.8	1.1	4.0	.44	36	1.1	46
8	1.9	.9	2.2	.58	35	2.1	22
9	2.7	1.0	4.1	.80	30	1.6	18
10	1.6	.5	4.2	.40	21	2.0	35
11	1.9	.5	3.1	.73	32	1.0	11
12	1.6	1.2	4.4	.54	45	1.6	23
13	.5	.3	1.7	.43	25	.2	10
14	4.0	1.5	3.1	.44	21	2.0	18
15	2.9	2.2	4.4	.23	20	.5	25
Median	*2.0*	*1.0*	*3.3*	*.44*	*28*	*1.2*	*23*
Ave.	*2.2*	*1.1*	*3.5*	*.54*	*30*	*1.3*	*25*
M.V.	*.7*	*.4*	*.8*	*.18*	*8.6*	*.52*	*7*

Statement of the results

Localization of sound.—In this test the blind are slightly superior to the seeing, especially in terms of the median. The blind vary more than the seeing, the coefficient of variability being 33 and 48 for the blind and the seeing respectively. The

records of the former range from 1.2 to 5.2 degrees while the latter vary from .5 to 4.4. The finest record of all was made by a seeing person.

The intensity of sound.—In this test the blind are slightly superior to the seeing according to the median but inferior by the same amount according to the average. The coefficient of variability is practically the same, namely, 32 for the blind and 33 for the seeing. Here again the best record of all is made by a seeing person, the same one, No. 13, who made the best record in localization of sound.

Lifted weights.—In this test the blind are inferior to the seeing both according to the median and the average. The coefficient of variability is approximately the same, 23 for the blind and 24 for the seeing. Here the finest record made is the same for the blind as for the seeing. But there are more extremely poor records among the blind than among the seeing.

Pressure.—In this test the blind are somewhat superior to the seeing both according to the median and the average. The blind are more variable, the coefficient being 40 for the blind and 33 for the seeing. The records of the blind range from .18 to 1.31 and those of the seeing from .13 to .98.

Touch.—In this test the larger the number the better the record. The facts therefore show that the blind are inferior to the seeing both according to the median and the average. There is but slight difference in the variability although the seeing have four records that are better than the best record of the blind.

Tactual space on the finger.—In this test there is no significant difference in the median and the average. But the variability of the seeing is greater, the coefficient being 40 for the seeing and 27 for the blind. The seeing have two records that are better than the best of the blind.

Tactual space on the arm.—Here again there is no significant difference in the median and the average nor in the coefficient of variability.

The general conclusion then so far as the net results of

the facts are concerned must be this, that, while there are slight differences shown in the Tables, there is no *significant* constant tendency in the records to favor either the blind or the seeing. In other words, *the blind and the seeing under these circumstances are, on the whole, equally sensitive to the direction of sound, intensity of sound, lifted weight, passive pressure, active pressure, and tactual space.*

Interpretation of these findings

These experiments were planned on the assumption that the tests would be elemental; *i.e.*, that the measurements were so simple and natural that training, experience, and special skill would not affect the results. In other words we aimed to test native capacity as distinguished from acquired ability.

From one point of view we may say, therefore, that the outcome of this series of experiments is to prove that the specific tests here used are elemental, because the records do not favor those who have acquired very marked skill and efficiency in these senses. That itself is an important feat in experimental procedure, because it sets up a new standard for psycho-physic tests. Records of psychological and psycho-pedagogical experiments for the past twenty-five years are full of examples showing how capacities like those here mentioned improve with age, intelligence, and practice. If the above conclusion is right we must re-submit those findings to determine to what extent the improvement found was due to the maladaptation of the experiment and the lack of skill of the experimenter rather than to characteristic change in the persons measured.

We cannot too strongly emphasize the necessity of reducing the procedure in an experiment of this kind to its simplest terms so that the record shall be of the actual capacity rather than of the possession of skill in adjusting oneself to a complex situation. There are scores of pitfalls in each one of the above experiments, which, if allowed to enter uncontrolled, would detract from the elemental nature of the test.

If we regard the results then from the converse side, they

prove that the extraordinary skill in guidance through hearing and pressure and other senses in the blind has its basis in apperception rather than in sensation or in the direct and immediate judgment of difference. The person who reads the point print has learned to give meaning to sensory impressions which the seeing person tends to neglect although possessed of the same power as the blind in sensing these differences.

If, then, we say that the blind person who reads point print, almost as we read with our eyes, and guides himself through hearing in many respects almost as we guide ourselves with our eyes, and we regard these performances as representative of the thousands and thousands of situations which occur in life, we arrive at the conclusion that the seeing have countless wonderful capacities lying latent and unrecognized so far as conscious use is concerned. The apparent squandering of sensory capacities may well be compared with the great waste in the struggle for existence by various forms of prolific extravagance in reproduction. We are very richly endowed with capacities of which we employ only relatively few, and these in a very inadequate way. The achievements of Helen Keller, *e.g.*, illustrate this.

But after all, it is difficult to appreciate the full extent to which we actually do depend upon the sort of fine sensitiveness here measured. While most of us do not read with our finger tips or guide ourselves with our ears when walking, we do make constant use of these senses in ordinary activities, most of it unconsciously. It may, therefore, be said, from one point of view that, that while seeing persons are not trained in certain specific acts of skill, they are trained in the use of the sensibilities involved in such skill. Indeed we are born with hierarchies of blind "instinctive skill". We often appeal from the sense of sight to touch and we constantly respond to the environment through hearing without resort to the physical mechanism in terms of which we think.

Technically all these tests involved discrimination as distinguished from sensitiveness or sensibility. The sensory judg-

ment is reduced to its lowest terms and is not far different from perception in its immediacy.

The conclusions of these experiments may then be summed up in the following statements:

First. The blind who are skillful in the use of touch, muscle sense, and hearing are not more sensitive or keen in sensory discrimination than seeing persons when fundamental capacities are tested.

Second. The necessity of making our psycho-physical tests elemental is strikingly indicated by the results of these tests.

Third. Development of the use of a sense consists not in the heightening of sensitivity or sensory discrimination but in the development of complexes and meanings in terms of these.

THE ELEMENTAL CHARACTER OF SENSORY DISCRIMINATION

BY

CARL E. SEASHORE AND KWEI TAN

The problem here reported grew out of that of the preceding article by Seashore and Ling in which it was demonstrated that the blind, though so largely guided by the sense of hearing and touch in their daily activities, are not superior to seeing persons in elemental tests of fundamental capacities in these senses. In that investigation an elemental test was defined as one in which the "measurements were so simple and natural that training, experience, and special skill would not affect the results." This fact of the equality of the blind and the seeing was regarded as proof of the elemental character of the test.

To gather proof from another angle it was decided to take one of these tests, namely, discrimination for the intensity of sound, and submit a number of observers to prolonged practice for the purpose of determining whether or not ability in this test improved with practice. The procedure was exactly the same as reported in the preceding article. The audiometer was used with the 100 v.d. fork, the current being kept at 1.2 volts. The mercury was kept clean, and the physical conditions and the control of the instrument were satisfactory throughout. A single increment was employed for the whole series, namely, the step from 38 to 39, on the audiometer, and the test consisted of one hundred trials for each day under conditions as nearly constant as possible, both subjective and objective. The record was kept in terms of the number of right cases. The observers, designated by letter in the following table, were all students in Psychology who had had some experience in experimental work but no training in this particular test.

In a few cases two or even three hundred trials were taken

during the same day, but the general effort was to get one set of observations a week. The experiments were conducted in the light-sound-and-jar-proof room in the psychological laboratory with both observer and experimenter present in the room. The room was ventilated freely every twenty minutes.

Records of these trials are shown in Table 1, in which the numbers given are the number of right decisions in each one hundred trials for each of the chosen observers. The average and median are given only for thirteen periods. It is evident that it would not be fair to continue the average beyond that point for the reason that each observer represented a different degree of ability and the dropping out of, *e.g.*, D. and H. W. at the end of the thirteenth day would distort the average record seriously.

The fluctuation in these figures may appear at first as being rather large, but when one considers that the record is in terms of right and wrong cases, there is really a very satisfactory uniformity on the whole.

Records of introspections and objective observations were kept faithfully, but as these are of an ordinary nature it is hardly worth space to detail them here. Suffice it to say that the variations in the per cent. of right cases is not greater than one would ordinarily find in a single one hundred trials on successive days under very favorable circumstances in any type of discrimination.

TABLE I—*The effect of practice on discrimination for the intensity of sound*

	A.	F.	H.	T.	G.W.	N.W.	C.	D.	Go.	Gr.	H.W.	*Median*	*Ave.*
1.	88	91	94	89	86	65	88	76	86	95	74	88.1	84.7
2.	90	92	92	89	88	64	95	75	88	94	65	89.0	84.7
3.	94	97	92	83	91	69	89	78	90	90	77	89.7	86.3
4.	93	95	80	78	91	87	97	86	80	93	79	87.3	87.1
5.	92	78	76	68	91	89	95	83	89	92	76	88.0	84.4
6.	91	80	73	72	87	89	90	68	94	88	82	86.5	83.0
7.	94	99	75	89	79	87	98	69	89	98	89	90.8	87.8
8.	92	99	90	94	88	91	95	82	94	96	76	91.5	90.6
9.	92	88	93	92	83	89	99	78	89	93	63	90.0	87.1
10.	88	90	95	95	94	85	88	82	82	86	69	87.0	86.7
11.	86	94	91	93	98	75	97	81	83	76	71	84.5	85.9
12.	87	91	78	85	100	79	90	82	87	91	61	86.0	84.6
13.	95	91	89	82	95	79	91	84	85	85	59	86.4	85.0
14.	87	96	84	93	95	90	89		93				
15.	89	96	99	85	94	75	90		80				
16.	88	96	94	96	91	84	90						
17.	90	93	96	95	85	90	93						
18.	83	86	94	91	94	86							
19.	95	96	97	94	96	93							
20.	84	96	96	67	92	90							

Looking then at the average and the median, we see that these run practically constant for the thirteen days. Beyond the thirteenth day each observer's record must be judged by itself, and it will be found that in these there is no marked tendency toward raising or lowering the general average for a given observer. We must, therefore, arrive at the conclusion that, for periods of training represented by this series, there is in this test no marked evidence of improvement with practice. For a single observer one hundred trials are inadequate, by the present method, but by taking a larger number of observers under similar conditions and by treating the results together in the average and the median, we compensate in part for this shortcoming.

Some of the fluctuations must, however, be accounted for by recognizing a certain degree of complexity in the problem. For example, various observers reported from time to time that they observed, or were trying, accessory aids in judging; for example, the sound was visualized and the judgment was made in terms of distance, the fainter sound being farther away. Or the difference was imaged in terms of motor tendencies, the lower sound calling forth, in one type, less effort because it was easy to hear and, in another type, the larger effort as a matter of counteraction. On the whole the introspections do not reveal any case in which the adoption of a particular method seems to have changed capacity permanently from the point of adoption.

Furthermore, the situation was not entirely free from objective disturbances. The observations being taken in the sound-light-and-jar-proof room, exceedingly faint accessory stimuli did prove disturbing. In an earlier experiment of this kind, one observer said that he could hear his backbone creak under these circumstances. This was probably not true, but it illustrates the tendency to be affected by obscure stimuli. There are, of course, momentary disturbances of itching, pain, or pressure, due to the holding of the receiver to the ear. Together with this there is also the subjective factor, the change

in disposition, physiological tone, aggressiveness, etc. from day to day.

Bearing in mind that this test was selected as a sample from among six other tests which were supposed to be elemental, it is reasonable to conclude, in the absence of evidence to the contrary, that the same elemental character could be established for the other tests reported in the foregoing article.

The only one of these on which there is any considerable amount of reliable literature available is that of lifted weights. This test has, perhaps, been studied with more exactness than any other test in the psycho-physical literature. Extraordinary refinements have been made in the discovery of the conditions for change in the capacity. Most of these are in the nature of the laws of illusions. But, as we look back upon literature as a whole, we find that ordinarily a careful observer does not improve with training. He may and usually does go through a cycle of changes which may be traced to the operation of certain motives for illusion. The same would probably be observed in any psycho-physical test that can be subjected to sufficiently rigorous study.

We have, then, in this and the preceding article evidence of the elemental character of the tests from two points of view, namely, that of highly specialized sense training in actual life, in the blind, and that of accurately controlled measurement of training in a comparatively short series.

There is a third line of evidence that must be sought. In fact, these experiments were suggested by the discovery that in pitch discrimination the actual native capacity in sensory discrimination for pitch does not develop with age. Whether that rule applies to these six tests should be worked out with great care upon a series of ages of children from, say eight to sixteen. On all the tests, we have records of children showing clearly that there is improvement in the record, which we may call the "gross record" or "total output", from year to year up to about fifteen. But the challenge set up by this experiment is to repeat those experiments and analyze them to find out to what extent this improvement with age consists

in overcoming the difficulties which are due to the incomplete mastery of the test by the observers. In other words, we may lay down this as an hypothesis for further experiment: In psycho-physic tests of sensitivity or discrimination the conditions must be so simple and so accurately under control that they are equally fair to the skilled and the unskilled, the trained and the untrained, the young and the old. Only then shall we measure sensory discrimination in its true capacity.